

Cartografia Stocastica e Teoria dei Processi Aleatori su Varietà

La cartografia classica tratta la superficie terrestre come un oggetto deterministico perfettamente conosciuto, dove ogni punto possiede coordinate precise e immutabili. Questa idealizzazione, sebbene matematicamente elegante e storicamente fruttuosa, si scontra inevitabilmente con la realtà dei processi di misurazione e rappresentazione geografica. Ogni osservazione della superficie terrestre è affetta da errori strumentali, ogni rilevamento topografico incorpora incertezze, ogni digitalizzazione introduce rumore. Inoltre, molti fenomeni geografici sono intrinsecamente aleatori: la morfologia costiera determinata da processi erosivi stocastici, la distribuzione della vegetazione influenzata da fattori climatici variabili, la struttura delle reti urbane emergente da innumerevoli decisioni individuali. La cartografia stocastica riconosce questa incertezza come caratteristica fondamentale e ineliminabile della rappresentazione geografica, sviluppando un apparato matematico rigoroso per modellarla, quantificarla e propagarla attraverso le trasformazioni cartografiche. Invece di considerare una mappa come una funzione deterministica $f: S^2 \rightarrow \mathbb{R}^2$ dalla sfera al piano, la concepiamo come una variabile aleatoria, ovvero una famiglia di mappe possibili equipaggiata con una distribuzione di probabilità che codifica la nostra conoscenza e la nostra incertezza. Questo approccio richiede l'estensione della teoria dei processi stocastici, tradizionalmente sviluppata in spazi euclidei, alle varietà riemanniane curve. La geometria intrinseca della sfera introduce complicazioni profonde: la covarianza spaziale deve rispettare le isometrie del gruppo $SO(3)$, le nozioni di stazionarietà devono essere ridefinite in assenza di traslazioni globali, i processi di diffusione seguono geodetiche anziché linee rette. Emerge così una teoria ricca e tecnicamente sofisticata che unifica probabilità, geometria differenziale e analisi funzionale.

Capitolo 1: Campi Aleatori Gaussiani su Varietà Riemanniane

1.1 Spazi di Probabilità su Varietà e Misure Gaussiane

Consideriamo una varietà riemanniana compatta (M, g) di dimensione n . Nel nostro contesto cartografico, M sarà tipicamente la sfera S^2 o un ellissoide di rotazione, ma la teoria che sviluppiamo mantiene validità generale. Uno spazio di probabilità sulla varietà è una tripla $(M, \mathcal{F}, \mu)$ dove $\mathcal{F}$ è la σ -algebra di Borel generata dalla topologia di M , e μ è una misura di probabilità su $\mathcal{F}$.

La misura più naturale su una varietà riemanniana è la misura di volume riemanniana dV_g , definita localmente in coordinate $(x^1, \dots, x^n)$ come:

$$dV_g = \sqrt{\det(g)} \, dx^1 \wedge \dots \wedge dx^n$$

dove g_{ij} sono le componenti del tensore metrico. Normalizzando questa misura, otteniamo la misura di probabilità uniforme:

$$\mu_{\text{unif}}(A) = \text{Vol}(A) / \text{Vol}(M)$$

per ogni insieme misurabile $A \subset M$. Sulla sfera S^2 di raggio R , questa misura coincide con la misura superficiale normalizzata:

$$\mu_{\text{unif}}(A) = \text{Area}(A) / (4\pi R^2)$$

Un campo aleatorio su M è una collezione di variabili aleatorie $\{X(p) : p \in M\}$ indicizzate dai punti della varietà, tutte definite sullo stesso spazio di probabilità astratto $(\Omega, \mathcal{A}, \mathbb{P})$. Formalmente, X è una funzione misurabile:

$$X: \Omega \times M \rightarrow \mathbb{R}$$

tale che per ogni $p \in M$ fissato, $\omega \mapsto X(\omega, p)$ è una variabile aleatoria, e per ogni $\omega \in \Omega$ fissato, $p \mapsto X(\omega, p)$ è una funzione (possibilmente irregolare) su M .

1.2 Definizione Rigorosa di Campo Gaussiano su Varietà

Un campo aleatorio X su M è detto gaussiano se per ogni scelta finita di punti $p_1, \dots, p_k \in M$, il vettore aleatorio $(X(p_1), \dots, X(p_k))$ ha distribuzione gaussiana multivariata in $\mathbb{R}^k$. Questa proprietà di coerenza delle distribuzioni finite-dimensionali è la caratterizzazione probabilistica fondamentale.

Un campo gaussiano è completamente determinato dalla sua funzione media $m: M \rightarrow \mathbb{R}$ e dalla sua funzione di covarianza $C: M \times M \rightarrow \mathbb{R}$, definite come:

$$m(p) = \mathbb{E}[X(p)] \quad C(p, q) = \mathbb{E}[(X(p) - m(p))(X(q) - m(q))]$$

La funzione di covarianza deve essere simmetrica: $C(p, q) = C(q, p)$, e definita positiva nel senso che per ogni scelta finita di punti $p_1, \dots, p_k$ e coefficienti reali $a_1, \dots, a_k$, deve valere:

$$\sum_{i=1}^k \sum_{j=1}^k a_i a_j C(p_i, p_j) \geq 0$$

Questa condizione garantisce che le distribuzioni gaussiane finite-dimensionali siano ben definite e consistenti.

Per molte applicazioni cartografiche, consideriamo campi a media nulla ($m(p) = 0$ per ogni p), concentrandoci sulla struttura di correlazione spaziale codificata dalla covarianza. La non-trivialità della teoria su varietà emerge già nella caratterizzazione delle covarianze ammissibili.

1.3 Il Teorema di Rappresentazione Spettrale su Varietà Compatte

Un risultato fondamentale della teoria dei campi aleatori su varietà compatte è il teorema di rappresentazione spettrale, che generalizza la decomposizione di Karhunen-Loève dagli spazi euclidei alle varietà curve. Questo teorema si basa sulla teoria spettrale dell'operatore di Laplace-Beltrami Δ_g .

L'operatore di Laplace-Beltrami su (M, g) è definito in coordinate locali come:

$$\Delta_g f = (1/\sqrt{\det(g)}) \partial_i (\sqrt{\det(g)} g^{ij} \partial_j f)$$

dove g^{ij} sono le componenti controvarianti del tensore metrico (l'inversa della matrice g_{ij}). Questo operatore è auto-aggiunto negativo rispetto al prodotto scalare $L^2(M, dV_g)$:

$$\langle f, \Delta_g h \rangle = \int_M f(p) \Delta_g h(p) dV_g(p)$$

Per varietà compatte, il teorema spettrale garantisce l'esistenza di una base ortonormale completa di $L^2(M)$ costituita da autofunzioni φ_k con autovalori λ_k :

$$\Delta_g \varphi_k = -\lambda_k \varphi_k$$

dove $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots$ con $\lambda_k \rightarrow \infty$. Le autofunzioni soddisfano:

$$\int_M \varphi_k(p) \varphi_j(p) dV_g(p) = \delta_{kj}$$

Il teorema di rappresentazione spettrale afferma che ogni campo gaussiano centrato X su M con covarianza regolare può essere rappresentato come:

$$X(p) = \sum_{k=0}^{\infty} \sqrt{\tilde{\lambda}_k} Z_k \varphi_k(p)$$

dove $\{Z_k\}$ sono variabili aleatorie gaussiane standard indipendenti ($Z_k \sim N(0,1)$) e i coefficienti $\tilde{\lambda}_k$ sono determinati dalla covarianza tramite:

$$C(p, q) = \sum_{k=0}^{\infty} \tilde{\lambda}_k \varphi_k(p) \varphi_k(q)$$

Questa rappresentazione è l'analogo della decomposizione in serie di Fourier per campi aleatori. La serie converge in $L^2(\Omega \times M)$ sotto condizioni di regolarità sulla covarianza.

1.4 La Funzione di Covarianza Isotropica sulla Sfera

Un caso particolarmente importante per la cartografia è quello dei campi isotropici sulla sfera S^2 . Un campo aleatorio X su S^2 è detto isotropico se la sua legge di distribuzione è invariante sotto l'azione del gruppo delle rotazioni $SO(3)$. In termini della covarianza, l'isotropia significa che $C(p, q)$ dipende solo dalla distanza geodetica $d(p, q)$ tra i punti:

$$C(p, q) = \tilde{C}(d(p, q))$$

per qualche funzione $\tilde{C}: [0, \pi R] \rightarrow \mathbb{R}$. La distanza geodetica sulla sfera di raggio R è data da:

$$d(p, q) = R \arccos(\langle p, q \rangle / R^2)$$

dove $\langle \cdot, \cdot \rangle$ denota il prodotto scalare euclideo sui vettori posizione in $\mathbb{R}^3$.

Equivalentemente, possiamo parametrizzare la covarianza isotropica mediante l'angolo $\theta = d(p, q)/R \in [0, \pi]$, scrivendo:

$$C(p, q) = \sigma^2 \rho(\theta)$$

dove $\sigma^2 = C(p, p)$ è la varianza (costante per l'isotropia) e $\rho: [0, \pi] \rightarrow \mathbb{R}$ è la funzione di correlazione normalizzata con $\rho(0) = 1$.

La caratterizzazione completa delle funzioni di correlazione ammissibili sulla sfera è fornita dal teorema di Schoenberg: una funzione continua $\rho: [0, \pi] \rightarrow \mathbb{R}$ con $\rho(0) = 1$ è una funzione di correlazione isotropica sulla S^2 se e solo se può essere espressa come:

$$\rho(\theta) = \sum_{\ell=0}^{\infty} a_{\ell} P_{\ell}(\cos \theta)$$

dove P_{ℓ} sono i polinomi di Legendre, i coefficienti a_{ℓ} sono non negativi, e la serie converge uniformemente su $[0, \pi]$. I coefficienti devono soddisfare:

$$\sum_{\ell=0}^{\infty} a_{\ell} (2\ell + 1) = 1$$

che deriva dalla normalizzazione $\rho(0) = 1$.

Questa rappresentazione è intimamente connessa alla teoria delle armoniche sferiche. Le autofunzioni del Laplaciano sulla sfera sono le armoniche sferiche Y_{ℓ}^m , organizzate in multipletti di dimensione $2\ell + 1$ per ogni $\ell \geq 0$, con autovalore $\lambda_{\ell} = \ell(\ell + 1)/R^2$. La decomposizione spettrale di un campo isotropico diventa:

$$X(p) = \sum_{\ell=0}^{\infty} \sqrt{a_{\ell} \sigma^2} \sum_{m=-\ell}^{\ell} Z_{\ell}^m Y_{\ell}^m(p)$$

dove $\{Z_{\ell}^m\}$ sono gaussiane standard indipendenti.

1.5 Esempi Fondamentali di Covarianze Sferiche

Presentiamo alcuni esempi importanti di funzioni di covarianza isotropiche sulla sfera, che emergono naturalmente in applicazioni cartografiche.

Covarianza esponenziale sferoide:

Una generalizzazione naturale della covarianza esponenziale euclidea $\rho(r) = \exp(-r/\ell)$ alla geometria sferica è:

$$\rho(\theta) = \exp(-2\sin(\theta/2)/\alpha)$$

dove $\alpha > 0$ è un parametro di scala angolare. Questa covarianza ha coefficienti spettrali:

$$a_{\ell} \sim (2\ell + 1) \exp(-\sqrt{\ell(\ell+1)}) \alpha \text{ per } \ell \rightarrow \infty$$

Il decadimento esponenziale in ℓ garantisce che i campi generati siano infinitamente differenziabili.

Covarianza di Matérn sferica:

La famiglia di Matérn sulla sfera generalizza la popolare famiglia euclidea. Per un parametro di smoothness $\nu > 0$ e scala angolare $\alpha > 0$:

$$\rho_{\nu}(\theta) = (2^{\nu} \Gamma(\nu)) (2\sin(\theta/2)/\alpha)^{\nu} K_{\nu}(2\sin(\theta/2)/\alpha)$$

dove K_v è la funzione di Bessel modificata del secondo tipo. I coefficienti spettrali si comportano come:

$$a_\ell \sim (2\ell + 1) \ell^{-2v-1} \text{ per } \ell \rightarrow \infty$$

Il parametro v controlla la regolarità: campi con $v > k$ sono k volte differenziabili quasi certamente. Per $v \rightarrow \infty$, si recupera la covarianza gaussiana.

Covarianza causale (damped spherical wave):

Questa covarianza emerge dalla soluzione dell'equazione d'onda smorzata sulla sfera:

$$\rho(\theta) = \exp(-\gamma\theta) \cos(\omega\theta + \varphi)$$

dove $\gamma > 0$ è il coefficiente di smorzamento, ω la frequenza angolare, e φ una fase. Questa famiglia permette di modellare correlazioni oscillatorie, utili per fenomeni periodici come pattern climatici stagionali.

1.6 Regolarità dei Percorsi e Teoria di Sobolev su Varietà

Una questione fondamentale nella teoria dei campi aleatori è la regolarità quasi certa dei percorsi campionari: con quale probabilità le realizzazioni $X(\cdot, \omega)$ sono funzioni continue, differenziabili, o di classe C^k ?

La risposta dipende dalla regolarità della covarianza, che a sua volta è codificata dal comportamento asintotico dei coefficienti spettrali a_ℓ . Definiamo gli spazi di Sobolev su M come:

$$H^s(M) = \{f \in L^2(M) : \sum_k \lambda_k^s |\langle f, \varphi_k \rangle|^2 < \infty\}$$

per $s \geq 0$, dove $\{\lambda_k, \varphi_k\}$ sono autovalori e autofunzioni di $-\Delta_g$. La norma di Sobolev è:

$$\|f\|_{H^s}^2 = \sum_k (1 + \lambda_k)^s |\langle f, \varphi_k \rangle|^2$$

Il teorema fondamentale di immersione di Sobolev su varietà compatte afferma che per $s > n/2 + k$ (dove $n = \dim M$), l'inclusione $H^s(M) \hookrightarrow C^k(M)$ è continua. Questo significa che funzioni con norma H^s finita sono automaticamente k volte continuamente differenziabili.

Per un campo gaussiano X con decomposizione spettrale $X = \sum_k \sqrt{\tilde{\lambda}_k} Z_k \varphi_k$, la condizione affinché $X \in H^s$ quasi certamente è:

$$\sum_k (1 + \lambda_k)^s \tilde{\lambda}_k < \infty$$

Sulla sfera S^2 , dove $\lambda_\ell = \ell(\ell + 1)$, con multiplicity $2\ell + 1$, questa diventa:

$$\sum_{\ell=0}^{\infty} (2\ell + 1)(1 + \ell(\ell+1))^s a_\ell < \infty$$

Teorema (Regolarità dei campi gaussiani su S^2): Sia X un campo gaussiano isotropico su S^2 con coefficienti spettrali a_ℓ . Se:

$$\sum_{\ell=0}^{\infty} (2\ell + 1) \ell^{2s} a_\ell < \infty$$

per qualche $s > 1 + k$, allora X ha percorsi C^k quasi certamente.

Per la covarianza di Matérn con parametro ν , la condizione diventa $\nu > 1 + k$, confermando che la regolarità del campo è controllata direttamente dal parametro di smoothness.

1.7 La Trasformata di Karhunen-Loève e Simulazione Efficiente

La decomposizione spettrale fornisce un metodo naturale per simulare realizzazioni di campi gaussiani su varietà. Data la rappresentazione:

$$X(p) = \sum_{k=0}^{\infty} \sqrt{\tilde{\lambda}_k} Z_k \varphi_k(p)$$

una realizzazione approssimata può essere ottenuta troncando la serie a un ordine finito K :

$$X^{(K)}(p) = \sum_{k=0}^K \sqrt{\tilde{\lambda}_k} Z_k \varphi_k(p)$$

dove Z_k sono estrazioni indipendenti da $N(0, 1)$. L'errore di troncamento in $L^2(M)$ ha varianza:

$$\mathbb{E}[\|X - X^{(K)}\|_{L^2}^2] = \sum_{k=K+1}^{\infty} \tilde{\lambda}_k$$

che decresce tanto più rapidamente quanto più velocemente decadono i coefficienti spettrali.

Sulla sfera S^2 , le autofunzioni sono le armoniche sferiche Y_{ℓ}^m . Per un campo isotropico, la realizzazione troncata al grado L è:

$$X^{(L)}(\theta, \varphi) = \sum_{\ell=0}^L \sqrt{a_{\ell} \sigma^2} \sum_{m=-\ell}^{\ell} Z_{\ell}^m Y_{\ell}^m(\theta, \varphi)$$

Il calcolo richiede:

1. Generazione di $(L+1)^2$ variabili gaussiane indipendenti Z_{ℓ}^m
2. Valutazione delle armoniche sferiche Y_{ℓ}^m nei punti desiderati
3. Sommazione pesata con $\sqrt{a_{\ell} \sigma^2}$

Le armoniche sferiche possono essere valutate efficientemente usando le formule ricorsive per i polinomi di Legendre associati:

$$Y_{\ell}^m(\theta, \varphi) = \sqrt{((2\ell+1)/(4\pi)) \cdot (\ell-m)!/(\ell+m)!} P_{\ell}^m(\cos \theta) e^{im\varphi}$$

dove P_{ℓ}^m sono i polinomi di Legendre associati, calcolabili via le ricorrenze di Bonnet.

Per grandi L , il costo computazionale è $O(L^4)$ usando algoritmi diretti, ma può essere ridotto a $O(L^2 \log^2 L)$ usando trasformate veloci sferiche basate sulla FFT non uniforme.

1.8 Anisotropia e Covarianze Non-Stazionarie

Molti fenomeni geografici reali esibiscono anisotropia: le correlazioni spaziali dipendono non solo dalla distanza ma anche dalla direzione. Sulla sfera, questo richiede covarianze della forma generale:

$$C(p, q) = \tilde{C}(p, q, n(p, q))$$

dove $n(p, q)$ è il vettore tangente unitario in p che punta verso q lungo la geodetica che li connette.

Una parametrizzazione utile sfrutta le coordinate geodetiche normali centrate in p . Dato $p \in S^2$, consideriamo la mappa esponenziale $\exp_p: T_p S^2 \rightarrow S^2$ che mappa vettori tangenti a punti sulla sfera via geodetiche. Per q vicino a p , esiste un unico vettore tangente $v \in T_p S^2$ tale che $q = \exp_p(v)$.

Una famiglia importante di covarianze anisotropiche sulla sfera è la famiglia Matérn geometrica:

$$C(p, q) = \sigma^2 (2^{1-\nu} \Gamma(\nu)) (\|\Sigma^{-1/2} \log_p(q)\|)^{2\nu} K_\nu(\|\Sigma^{-1/2} \log_p(q)\|)$$

dove $\log_p: S^2 \rightarrow T_p S^2$ è la mappa logaritmica (inversa dell'esponenziale), Σ è una matrice 2×2 definita positiva che codifica l'anisotropia in $T_p S^2$, e la norma è quella euclidea nel piano tangente.

Questa costruzione estende la covarianza di Matérn euclidea alla geometria sferica in modo geometricamente naturale. La matrice Σ determina gli assi principali di correlazione: autovalori diversi indicano correlazioni diverse lungo direzioni diverse.

Per covarianze non-stazionarie, Σ può variare spazialmente: $\Sigma = \Sigma(p)$. Questo genera campi con struttura di correlazione localmente variabile, essenziale per modellare fenomeni come topografia con orientazione variabile (catene montuose con diverse direzioni di allineamento in regioni diverse).

Capitolo 2: Processi di Wiener e Diffusioni su Varietà Curve

2.1 Il Moto Browniano Riemanniano: Definizione e Costruzione

Il moto browniano su una varietà riemanniana (M, g) è il processo stocastico che generalizza il moto browniano euclideo rispettando la geometria intrinseca di M . Esistono diverse caratterizzazioni equivalenti, ciascuna illuminante da prospettive diverse.

Definizione (via generatore infinitesimale): Il moto browniano su (M, g) è il processo di Markov $\{X_t\}_{t \geq 0}$ a valori in M il cui generatore infinitesimale è $(1/2)\Delta_g$, dove Δ_g è il Laplaciano di Laplace-Beltrami. Questo significa che per ogni funzione $f \in C^2(M)$:

$$\lim_{t \rightarrow 0} \frac{\mathbb{E}_p[f(X_t) - f(p)]}{t} = (1/2)\Delta_g f(p)$$

dove $\mathbb{E}_p$ denota l'aspettazione condizionale dato $X_0 = p$.

Definizione (via equazione differenziale stocastica): In coordinate locali $(x^1, \dots, x^n)$, il moto browniano è soluzione dell'equazione differenziale stocastica di Stratonovich:

$$dx^i = \sum_{j,k} \Gamma^i_{jk}(x) dx^j \circ dx^k + \sqrt{2} dB^i$$

dove Γ^i_{jk} sono i simboli di Christoffel della connessione di Levi-Civita, $\{B^i\}$ sono n moti browniani standard indipendenti, e $\circ$ denota l'integrale di Stratonovich.

Il fattore $\sqrt{2}$ nella diffusione deriva dalla normalizzazione del generatore come $(1/2)\Delta_g$. La forma di Stratonovich è geometricamente naturale: trasformazioni di coordinate agiscono come regola della catena ordinaria.

Definizione (via sviluppo stocastico): Globalmente, il moto browniano può essere costruito mediante sviluppo stocastico (stochastic development). Dato un moto browniano B_t in $\mathbb{R}^n$, lo "eleviamo" al fibrato dei frame ortonormali $O(M)$ (il fibrato principale con fibre costituite da frame ortonormali negli spazi tangenti), ottenendo un processo orizzontale nel fibrato, e poi proiettiamo su M . Questa costruzione, dovuta a Eells-Elworthy-Malliavin, è intrinseca e indipendente dalle coordinate.

2.2 Il Nucleo di Transizione e l'Equazione del Calore

Il moto browniano su M determina un semigruppato di operatori lineari $\{P_t\}_{t \geq 0}$ su funzioni misurabili limitate, definito da:

$$P_t f(p) = \mathbb{E}_p[f(X_t)]$$

Questo è il semigruppato di transizione del processo. Per varietà compatte, P_t ammette un nucleo integrale $p_t(x, y)$, chiamato nucleo del calore, tale che:

$$P_t f(x) = \int_M p_t(x, y) f(y) dV_g(y)$$

Il nucleo del calore è la soluzione fondamentale dell'equazione del calore:

$$\partial u / \partial t = (1/2) \Delta_g u$$

con condizione iniziale $u(0, x) = \delta_x$ (la delta di Dirac in x). Esplicitamente:

$$(\partial / \partial t - (1/2) \Delta_g) p_t(x, y) = 0 \quad \lim_{t \rightarrow 0} p_t(x, \cdot) = \delta_x$$

Il nucleo del calore codifica la densità di probabilità di transizione del moto browniano: $p_t(x, y)$ è la densità della probabilità che una particella browniana partente da x al tempo 0 si trovi in un intorno infinitesimale di y al tempo t .

2.3 Sviluppo Asintotico del Nucleo del Calore

Per tempi piccoli, il nucleo del calore ha uno sviluppo asintotico universale che codifica invarianti geometrici locali. Sulla sfera S^2 di raggio R :

$$p_t(p, q) \sim (1/(4\pi t)) \exp(-d^2(p, q)/(2t)) [1 + (R^2/(12t)) + O(t)]$$

dove $d(p, q)$ è la distanza geodetica. Il termine dominante $(1/(4\pi t)) \exp(-d^2/(2t))$ è gaussiano, riflettendo il comportamento localmente euclideo. La correzione di curvatura $R^2/(12t)$ appare al secondo ordine.

In generale, per una varietà n -dimensionale con curvatura scalare $S(x)$:

$$p_t(x, y) \sim (1/(4\pi t)^{n/2}) \exp(-d^2(x, y)/(2t)) [1 - (t/6)S(x) + O(t^2)]$$

Questo è il primo termine dell'espansione di Minakshisundaram-Pleijel, che può essere estesa a tutti gli ordini:

$$p_t(x, y) \sim (1/(4\pi t)^{n/2}) \exp(-d^2(x, y)/(2t)) \sum_{k=0}^{\infty} a_k(x, y) t^k$$

I coefficienti a_k sono determinabili algebricamente da invarianti curvaturali locali: $a_0 = 1$, a_1 coinvolge la curvatura scalare, a_2 coinvolge potenze seconde della curvatura e il tensore di Ricci, e così via. Il calcolo esplicito di questi coefficienti è un problema tecnico significativo in geometria spettrale.

2.4 La Rappresentazione Spettrale del Nucleo del Calore

Per varietà compatte, il nucleo del calore ammette una rappresentazione spettrale completa in termini degli autovalori e autofunzioni del Laplaciano:

$$p_t(x, y) = \sum_{k=0}^{\infty} e^{-\lambda_k t/2} \varphi_k(x) \varphi_k(y)$$

Questa serie converge uniformemente per ogni $t > 0$. Per $t \rightarrow \infty$, il nucleo converge alla distribuzione stazionaria:

$$\lim_{t \rightarrow \infty} p_t(x, y) = \varphi_0(x) \varphi_0(y) = 1/(\text{Vol}(M))$$

dove abbiamo usato che $\lambda_0 = 0$ e $\varphi_0 = 1/\sqrt{\text{Vol}(M)}$ è la funzione costante normalizzata.

Sulla sfera S^2 , la rappresentazione spettrale diventa:

$$p_t(p, q) = \sum_{\ell=0}^{\infty} (2\ell+1)/(4\pi R^2) \exp(-\ell(\ell+1)t/(2R^2)) P_\ell(\cos \theta)$$

dove θ è l'angolo geodetico tra p e q , e P_ℓ sono i polinomi di Legendre. Il fattore $(2\ell+1)/(4\pi R^2)$ deriva dalla degenerazione delle armoniche sferiche di grado ℓ .

Questa formula fornisce un metodo alternativo per simulare il moto browniano sferico: dato $X_0 = p_0$, la distribuzione di X_t è data dal nucleo $p_t(p_0, \cdot)$. Campionando da questa distribuzione mediante rejection sampling o MCMC, si ottiene una traiettoria del processo.

2.5 Coupling e Confronto di Processi di Diffusione

Una tecnica potente per studiare processi stocastici su varietà è il coupling: costruire due processi sullo stesso spazio di probabilità in modo che le loro traiettorie siano correlate in modo controllabile. Per varietà curve, i coupling sono intimamente legati alla curvatura.

Coupling di Kendall: Dati due moti browniani X_t e Y_t su (M, g) partenti da x_0 e y_0 , il coupling di Kendall li fa evolvere indipendentemente fino a quando si incontrano ($X_\tau = Y_\tau$ per qualche tempo aleatorio τ), dopo di che evolvono identicamente ($X_t = Y_t$ per $t \geq \tau$).

Il tempo di accoppiamento τ ha distribuzione che dipende criticamente dalla curvatura. Per varietà con curvatura di Ricci positiva (come la sfera), il coupling avviene rapidamente; per curvature negative, richiede tempi più lunghi.

Coupling mirror: Su varietà con simmetrie, si può costruire il coupling specchio: se X_t e Y_t sono equidistanti da un punto o da una sottovarietà totalmente geodetica Σ , li si fa evolvere simmetricamente rispetto a Σ . Questo è particolarmente utile sulla sfera, dove ogni coppia di punti antipodali determina naturalmente un equatore di simmetria.

Lemma di confronto delle distanze: Siano X_t e Y_t due moti browniani accoppiati su (M, g) con curvatura di Ricci $\text{Ric} \geq \kappa g$ per qualche $\kappa \in \mathbb{R}$. Allora la distanza $d_t = d(X_t, Y_t)$ soddisfa:

$$d(\mathbb{E}[d_t^2])/dt \leq 2n - \kappa \mathbb{E}[d_t^2]$$

dove $n = \dim(M)$. Per $\kappa > 0$ (curvatura positiva), questo implica che la distanza attesa decresce esponenzialmente:

$$\mathbb{E}[d_t^2] \leq e^{-\kappa t} d_0^2$$

Su S^2 con $R = 1$, la curvatura di Ricci è $\text{Ric} = g$, quindi $\kappa = 1$, e due moti browniani si avvicinano esponenzialmente con rate 1. Per l'ellissoide terrestre con semi-assi $a \approx 6378$ km e $b \approx 6357$ km, la curvatura varia leggermente, ma rimane positiva ovunque, garantendo mixing rapido.

2.6 Il Laplaciano Stocastico e Integrazione Stocastica su Varietà

Per definire integrali stocastici su varietà, dobbiamo generalizzare la costruzione di Itô dal caso euclideo. La difficoltà principale è che non esiste una struttura vettoriale globale su una varietà curva: non possiamo sommare punti o moltiplicarli per scalari in modo globalmente consistente.

La soluzione è lavorare negli spazi tangenti, che sono spazi vettoriali. Dato un processo X_t su M e un campo vettoriale V lungo la traiettoria (ossia $V_t \in T_{\{X_t\}M}$), vogliamo definire l'integrale stocastico:

$$\int_0^t V_s \cdot dX_s$$

La chiave è il trasporto parallelo stocastico. Data una traiettoria X_t , il trasporto parallelo $P_{s \rightarrow t}: T_{\{X_s\}M} \rightarrow T_{\{X_t\}M}$ lungo X tra i tempi s e t è definito dall'equazione:

$$\nabla_{\{dX/dt\}} V = 0$$

dove ∇ è la derivata covariante associata alla connessione di Levi-Civita. Il trasporto parallelo preserva il prodotto scalare: $\langle P_{s \rightarrow t}(v), P_{s \rightarrow t}(w) \rangle_{\{X_t\}} = \langle v, w \rangle_{\{X_s\}}$.

Definizione (integrale stocastico di Itô su varietà): Sia V_t un processo adattato a valori in $T_{\{X_t\}M}$. L'integrale stocastico è definito come:

$$\int_0^t V_s \cdot dX_s = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \langle P_{\{t_i \rightarrow t\}}(V_{\{t_i\}}), X_t - X_{\{t_i\}} \rangle$$

dove il limite è in probabilità su partizioni $t_0 = 0 < t_1 < \dots < t_n = t$ con mesh tendente a zero.

Questa definizione estende naturalmente la costruzione euclidea. Il trasporto parallelo "identifica" i diversi spazi tangenti, permettendo di sommare i contributi infinitesimali.

La formula di Itô su varietà assume una forma modificata che include termini di correzione curvaturale. Per un processo X_t soluzione di:

$$dX_t = A(X_t) dt + B(X_t) dB_t$$

dove A è un campo vettoriale e $B: M \rightarrow T^*M \otimes \mathbb{R}^k$, e data $f \in C^2(M)$, si ha:

$$df(X_t) = (A^i \partial_i f + (1/2) B^i_{\alpha} B^j_{\alpha} (\nabla_i \partial_j f)) dt + (B^i_{\alpha} \partial_i f) dB^{\alpha}_t$$

Il termine $\nabla_i \partial_j f$ contiene i simboli di Christoffel e codifica l'effetto della curvatura sulla variazione quadratica.

2.7 Processi di Diffusione Generali e Operatori Ellittici

Un processo di diffusione generale su M con drift b e diffusione σ è un processo di Markov X_t il cui generatore infinitesimale è:

$$\mathcal{L}f = b^i \partial_i f + (1/2) a^{ij} \nabla_i \partial_j f$$

dove $a = \sigma \sigma^T$ è la matrice di diffusione (un campo tensoriale simmetrico definito positivo su M), e b è un campo vettoriale.

L'equazione differenziale stocastica corrispondente in forma di Stratonovich è:

$$dx^i = b^i(x) dt + \sigma^i_{\alpha}(x) \circ dB^{\alpha}$$

In coordinate locali, la forma di Itô include correzioni:

$$dx^i = (b^i - (1/2) \sigma^j_{\alpha} \partial_j \sigma^i_{\alpha}) dt + \sigma^i_{\alpha} dB^{\alpha}$$

dove i termini aggiuntivi derivano dalla conversione Stratonovich-Itô.

Esistenza e unicità: Sotto condizioni di Lipschitz su b e σ , esiste una soluzione unica (forte) all'equazione fino al primo tempo di uscita da ogni compatto. Per varietà compatte, la soluzione è globale in tempo.

Invarianza della misura: Una misura μ su M è invariante per il processo se:

$$\int_M (\mathcal{L}f) d\mu = 0$$

per ogni $f \in C^2(M)$. Questo significa che μ è stazionaria: se $X_0 \sim \mu$, allora $X_t \sim \mu$ per ogni t .

Per il moto browniano standard ($b = 0$, $a = g$), la misura invariante è la misura di volume riemanniana normalizzata. Per diffusioni con drift, la misura invariante (quando esiste) ha densità proporzionale a:

$$\rho(x) \propto \exp(\int^x (a^{-1}b)^i dx_i)$$

rispetto alla misura di volume, dove l'integrale dipende dal percorso in generale (è ben definito quando esiste un potenziale).

2.8 Processi di Ornstein-Uhlenbeck Geometrici

Un caso particolarmente importante per applicazioni cartografiche è il processo di Ornstein-Uhlenbeck su varietà, che generalizza il processo euclideo:

$$dX_t = -\nabla V(X_t) dt + \sqrt{2T} dB_t$$

dove $V: M \rightarrow \mathbb{R}$ è una funzione potenziale e $T > 0$ è una "temperatura". Questo processo modella una particella soggetta a una forza restauratrice $-\nabla V$ e a rumore termico.

Il generatore è:

$$\mathcal{L} = -\langle \nabla V, \nabla \cdot \rangle + T \Delta_g$$

La misura di Gibbs:

$$\mu(dx) = Z^{-1} \exp(-V(x)/T) dV_g(x)$$

dove $Z = \int_M \exp(-V/T) dV_g$ è la funzione di partizione, è invariante e inoltre reversibile: il processo soddisfa il bilancio dettagliato:

$$p_t(x, y) \mu(dx) = p_t(y, x) \mu(dy)$$

Questa proprietà è fondamentale per algoritmi MCMC (Markov Chain Monte Carlo) su varietà.

Convergenza all'equilibrio: Se V è strettamente convessa (nel senso che $\text{Hess}V \geq \lambda g$ per qualche $\lambda > 0$), il processo converge esponenzialmente alla misura di Gibbs:

$$\|p_t(x, \cdot) - \mu\|_{TV} \leq C e^{-\lambda t}$$

dove $\|\cdot\|_{TV}$ è la distanza in variazione totale. Il rate λ è controllato dalla costante di convessità di V e dalla geometria di M (curvatura, diametro).

Applicazione cartografica: Per modellare l'elevazione topografica con vincoli di smoothness, si può usare un processo di Ornstein-Uhlenbeck con potenziale $V(h) = \alpha \int_M \|\nabla h\|^2 dV_g$ che penalizza gradients elevati. Le realizzazioni campionarie sono funzioni smooth sulla sfera che bilanciano fedeltà ai dati e regolarità geometrica.

Capitolo 3: Filtri di Kalman Geometrici per Aggiornamento Bayesiano

3.1 Il Problema del Filtraggio su Varietà

Il problema del filtraggio stocastico consiste nell'inferire lo stato X_t di un sistema dinamico a partire da osservazioni parziali e rumorose Y_s per $s \leq t$. Quando lo stato vive su una varietà M anziché in uno spazio euclideo, emergono sfide tecniche significative che richiedono l'estensione della teoria classica.

Consideriamo il modello generale:

Equazione dello stato: $dX_t = f(X_t, t) dt + G(X_t, t) dW_t$

dove $X_t \in M$, $f: M \times \mathbb{R}_+ \rightarrow TM$ è il drift, $G: M \times \mathbb{R}_+ \rightarrow TM \otimes \mathbb{R}^m$ è la diffusione, e W_t è un moto browniano m -dimensionale standard.

Equazione dell'osservazione: $Y_t = h(X_t, t) + V_t$

dove $Y_t \in \mathbb{R}^p$ è l'osservazione, $h: M \times \mathbb{R}_+ \rightarrow \mathbb{R}^p$ è la funzione di misura, e V_t è rumore gaussiano bianco con covarianza $R(t)$.

L'obiettivo è calcolare la distribuzione condizionale posteriore $\pi_t = \mathcal{L}(X_t | Y_s, s \leq t)$, che codifica tutta l'informazione disponibile su X_t date le osservazioni fino al tempo t .

3.2 L'Equazione di Kushner-Stratonovich su Varietà

La distribuzione posteriore π_t evolve secondo l'equazione di Kushner-Stratonovich, che è la versione stocastica dell'equazione di Fokker-Planck. Per una funzione test $\varphi: M \rightarrow \mathbb{R}$, l'evoluzione dell'aspettazione posteriore $\mathbb{E}^\pi \pi_t[\varphi] = \int_M \varphi(x) \pi_t(dx)$ è:

$$d\mathbb{E}^\pi \pi_t[\varphi] = \mathbb{E}^\pi \pi_t[\mathcal{L}\varphi] dt + (\mathbb{E}^\pi \pi_t[\varphi h] - \mathbb{E}^\pi \pi_t[\varphi] \mathbb{E}^\pi \pi_t[h])^T R^{-1} (dY_t - \mathbb{E}^\pi \pi_t[h] dt)$$

dove $\mathcal{L}$ è il generatore infinitesimale del processo X_t senza osservazioni:

$$\mathcal{L}\varphi = \langle f, \nabla \varphi \rangle + (1/2) \text{Tr}(GG^T \text{Hess } \varphi)$$

Il termine $dY_t - \mathbb{E}^\pi \pi_t[h] dt$ è l'innovazione: la differenza tra l'osservazione effettiva e quella predetta dalla distribuzione corrente.

Questa equazione è esatta ma infinite-dimensionale: descrive l'evoluzione di una distribuzione di probabilità su M , che è un oggetto funzionale di dimensione infinita. Per implementazioni pratiche, si richiedono approssimazioni finite-dimensionali.

3.3 Il Filtro di Kalman Esteso su Varietà (EKF-M)

Quando la varietà M è uno spazio euclideo, e i modelli sono lineari gaussiani, il filtro di Kalman fornisce formule ricorsive esatte per la media e covarianza posteriori. Su varietà curve, la linearizzazione locale porta al filtro di Kalman esteso geometrico.

L'idea è rappresentare la distribuzione posteriore mediante la sua approssimazione gaussiana. Sulla varietà, una "distribuzione gaussiana" centrata in $m \in M$ con covarianza P è definita via logarithmic map:

$$\pi_t \approx \mathcal{N}_M(m_t, P_t)$$

dove per x vicino a m :

$$\pi_t(dx) \propto \exp(-(1/2)(\log_m(x), P^{-1} \log_m(x))) dV_g(x)$$

e $\log_m: U \rightarrow T_m M$ è la mappa logaritmica (inversa dell'esponenziale) definita in un intorno U di m .

Passo di predizione: Data la distribuzione a priori $\mathcal{N}_M(m_t, P_t)$ al tempo t (appena prima dell'osservazione), propaghiamo attraverso la dinamica fino a t^+ :

$$\tilde{m}_t = \exp_{\{m_t\}}(\int_{\{t\}^{t^+}} f(m_s, s) ds)$$

$$\tilde{P}_t = F_t P_t F_t^T + Q_t$$

dove $F_t = D\exp_{\{m_t\}}(\int f ds)$ è la matrice Jacobiana dell'esponenziale (derivata della mappa esponenziale), e Q_t è la covarianza del rumore di processo.

Passo di aggiornamento: All'arrivo di una nuova osservazione y_t , aggiorniamo:

$$K_t = \tilde{P}_t H_t^T (H_t \tilde{P}_t H_t^T + R_t)^{-1}$$

$$m_{t^+} = \exp_{\{\tilde{m}_t\}}(K_t (y_t - h(\tilde{m}_t, t)))$$

$$P_{t^+} = (I - K_t H_t) \tilde{P}_t$$

dove $H_t = \nabla h(\tilde{m}_t, t): T_{\{\tilde{m}_t\}}M \rightarrow \mathbb{R}^p$ è il differenziale della funzione di osservazione, e K_t è il guadagno di Kalman.

La differenza chiave rispetto all'EKF euclideo è l'uso delle mappe esponenziale e logaritmica per "trasportare" correzioni dallo spazio tangente alla varietà e viceversa.

3.4 Il Filtro di Kalman Unscented su Varietà (UKF-M)

L'EKF-M si basa sulla linearizzazione, che può essere imprecisa per non-linearità forti o curvatura elevata. Il filtro unscented propaga l'incertezza attraverso un insieme deterministico di sigma points scelti in modo da catturare media e covarianza.

Generazione dei sigma points: Data una distribuzione $\mathcal{N}_M(m, P)$ su M , generiamo $2n + 1$ punti (dove $n = \dim M$):

$$\chi_0 = m \quad \chi_i = \exp_m(\sqrt{((n + \lambda) P)} \cdot e_i) \text{ per } i = 1, \dots, n \quad \chi_{i+n} = \exp_m(-\sqrt{((n + \lambda) P)} \cdot e_i) \text{ per } i = 1, \dots, n$$

dove e_i sono i vettori della base ortonormale in $T_m M$, λ è un parametro di scaling, e $\sqrt{P}$ denota la radice quadrata matriciale (via decomposizione di Cholesky).

Trasformazione unscented: Per propagare attraverso una funzione $g: M \rightarrow N$:

1. Trasformiamo i sigma points: $\tilde{y}_i = g(\chi_i)$
2. Calcoliamo la media pesata su N : $\tilde{y} = \sum_i w_i \tilde{y}_i$ (usando la media intrinseca di Fréchet)
3. Calcoliamo la covarianza: $P^y = \sum_i w_i \log_{\tilde{y}}(\tilde{y}_i) \otimes \log_{\tilde{y}}(\tilde{y}_i)$

dove w_i sono pesi opportuni ($w_0 = \lambda/(n + \lambda)$, $w_i = 1/(2(n + \lambda))$ per $i \geq 1$).

Questo approccio non richiede calcolo di Jacobiani e cattura approssimazioni di ordine superiore della distribuzione trasformata.

Applicazione cartografica: Per tracciare la posizione di un veicolo sulla superficie terrestre con GPS rumoroso, lo stato è $X_t \in S^2$ (posizione sulla sfera). Le osservazioni sono coordinate lat-lon corrotte da rumore gaussiano. L'UKF-M propaga l'incertezza rispettando la geometria sferica, evitando artefatti che emergerebbero trattando (lat, lon) come coordinate cartesiane.

3.5 Filtri Particellari su Varietà

Per non-linearità severe o distribuzioni posteriori multi-modali, i filtri gaussiani (EKF, UKF) sono inadeguati. I filtri particellari rappresentano la distribuzione posteriore mediante un insieme discreto di campioni (particelle) con pesi.

La distribuzione posteriore è approssimata da:

$$\pi_t \approx \pi_t^N = \sum_{i=1}^N w_i^t \delta_{\{X_i^t\}}$$

dove $\{X_i^t\}_{i=1}^N$ sono particelle su M , $w_i^t \geq 0$ sono pesi con $\sum_i w_i = 1$, e δ_x è la delta di Dirac in x .

Algoritmo del filtro particellare su M :

1. **Inizializzazione ($t = 0$):** Campiona N particelle dalla distribuzione iniziale: $X_i^0 \sim \pi_0$, $w_i^0 = 1/N$
2. **Predizione ($t^- \rightarrow t$):** Propaga ogni particella attraverso la dinamica:
 - Campiona rumore $W_i \sim \mathcal{N}(0, I_m)$
 - Risolvi l'SDE: $dX_{i,s} = f(X_{i,s}, s) ds + G(X_{i,s}, s) dW_{i,s}$ per $s \in [t^-, t]$
3. **Aggiornamento (all'osservazione y_t):** Ricalcola i pesi:
 - $\tilde{w}_i^t = w_i^{t^-} \cdot p(y_t | X_i^t) = w_i^{t^-} \cdot \mathcal{N}(y_t; h(X_i^t, t), R)$
 - Normalizza: $w_i^t = \tilde{w}_i^t / \sum \tilde{w}_i^t$
4. **Resampling:** Se $ESS = 1/\sum (w_i^t)^2 < N_{\text{thresh}}$, ricampiona particelle con probabilità proporzionali ai pesi, reimpostando $w_i = 1/N$.

Il resampling previene la degenerazione: senza di esso, dopo poche iterazioni tutti i pesi si concentrerebbero su una singola particella.

Stimatori: La stima posteriore è:

$$\mathbb{E}^{\pi_t}[\varphi] \approx \sum_{i=1}^N w_i^t \varphi(X_i^t)$$

Per la media sulla varietà, si usa la media di Fréchet (o Riemanniana):

$$m_t = \operatorname{argmin}_{\{m \in M\}} \sum_{i=1}^N w_i^t d^2(m, X_i^t)$$

che generalizza il baricentro pesato rispettando la geometria di M .

3.6 La Divergenza di Kullback-Leibler su Varietà e Information Geometry

Per quantificare quanto un'approssimazione del filtro discosta dalla distribuzione posteriore vera, si usa la divergenza di Kullback-Leibler (KL):

$$D_{\{KL\}}(\pi \parallel \tilde{\pi}) = \int_M \log(d\pi/d\tilde{\pi}) d\pi$$

dove π è la distribuzione vera e $\tilde{\pi}$ l'approssimazione. Su varietà, questa quantità ha interpretazioni geometriche profonde legate alla geometria dell'informazione di Amari.

Lo spazio $\mathcal{A}(M)$ delle distribuzioni di probabilità su M può essere dotato di una struttura riemanniana mediante la metrica di Fisher:

$$g^{\{Fisher\}}_{\{\pi\}}(V, W) = \int_M (\partial_V \log \pi)(\partial_W \log \pi) \pi(dx)$$

dove V, W sono "direzioni" tangenti a $\mathcal{A}(M)$ (variazioni infinitesimali di π).

Le geodetiche rispetto alla metrica di Fisher sono le famiglie esponenziali. La divergenza KL è asintoticamente quadratica lungo queste geodetiche:

$$D_{\{KL\}}(\pi_0 \parallel \pi_t) \approx (t^2/2) g^{\{Fisher\}}(V, V) + O(t^3)$$

dove $\pi_t = \exp(t V)\pi_0$ è una famiglia esponenziale con direzione V .

Teorema (Information inequality): Per qualsiasi stimatore $\hat{m}_t$ basato su osservazioni fino a t , l'errore quadratico medio soddisfa:

$$\mathbb{E}[(d(\hat{m}_t, X_t))^2] \geq (I_t)^{-1}$$

dove I_t è l'informazione di Fisher accumulata fino al tempo t . Questa è la versione geometrica del bound di Cramér-Rao, che fornisce un limite inferiore fondamentale sulla precisione stimabile.

Per il filtro di Kalman su varietà, quando il modello è "localmente lineare gaussiano", l'EKF-M raggiunge asintoticamente questo bound, dimostrando la sua efficienza asintotica.

3.7 Applicazioni: Tracking su Sfera e Aggiornamento di Mappe Topografiche

Esempio 1: Tracking di un satellite in orbita

Lo stato del satellite è (posizione, velocità) su TS^2 (il fibrato tangente della sfera). La dinamica nel riferimento inerziale è:

$$d/dt(r, v) = (v, -\mu r/|r|^3 + \text{perturbazioni})$$

dove μ è il parametro gravitazionale terrestre. Le osservazioni sono misure GPS rumorose della posizione.

Implementando un UKF-M, i sigma points catturano la curvatura della traiettoria orbitale. La predizione propaga i punti seguendo la dinamica kepleriana, l'aggiornamento corregge usando le osservazioni GPS. La covarianza posteriore codifica l'ellissoide di incertezza sulla posizione.

Esempio 2: Aggiornamento sequenziale di DTM (Digital Terrain Model)

Un DTM è una funzione $h: M \rightarrow \mathbb{R}$ che assegna l'elevazione ad ogni punto sulla superficie di riferimento M . Modellando h come realizzazione di un campo gaussiano con covarianza Matérn, e ricevendo misure puntuali rumorose $\tilde{h}(p_i) = h(p_i) + \epsilon_i$, il problema è inferire h dato $\{(p_i, \tilde{h}(p_i))\}$.

Questo è un problema di filtraggio infinite-dimensionale: lo stato $X_t = h$ è un campo, le osservazioni sono funzionali lineari $Y_t = h(p_t) + V_t$. La distribuzione posteriore è una misura gaussiana sullo spazio di funzioni.

Usando rappresentazioni finite-dimensionali (truncated Karhunen-Loève), riduciamo a un filtro di Kalman standard sui coefficienti spettrali. Alternativamente, metodi kernel diretti (kriging) forniscono la soluzione esatta in forma chiusa:

$$h(p) \mid \text{dati} \sim \mathcal{N}(\mu^{\{\text{post}\}}(p), C^{\{\text{post}\}}(p, q))$$

$$\mu^{\{\text{post}\}}(p) = C(p, \{p_i\}) [C(\{p_i\}, \{p_i\}) + \sigma^2 I]^{-1} \{\tilde{h}(p_i)\}$$

$$C^{\{\text{post}\}}(p, q) = C(p, q) - C(p, \{p_i\}) [C(\{p_i\}, \{p_i\}) + \sigma^2 I]^{-1} C(\{p_i\}, q)$$

dove $C(\{p_i\}, \{p_i\})$ è la matrice di covarianza valutata nei punti osservati.

Capitolo 4: Processi di Lévy e Salti su Superfici Curve

4.1 Introduzione ai Processi con Discontinuità

Molti fenomeni geografici esibiscono discontinuità intrinseche: faglie tettoniche, coste, confini politici, transizioni brusche di vegetazione. I processi di diffusione continui come il moto browniano non catturano questi salti. I processi di Lévy generalizzano permettendo traiettorie discontinue.

Un processo di Lévy X_t su $\mathbb{R}^n$ è caratterizzato dalla proprietà che gli incrementi $X_{t+s} - X_t$ sono stazionari e indipendenti, con distribuzione dipendente solo da s . La legge infinitesimale è determinata dalla tripla di Lévy (γ, Σ, ν) :

- $\gamma \in \mathbb{R}^n$: drift
- Σ : matrice di covarianza della componente browniana
- ν : misura di Lévy sui salti, che soddisfa $\int \min(|x|^2, 1) \nu(dx) < \infty$

La formula di Lévy-Khintchine esprime la funzione caratteristica:

$$\mathbb{E}[e^{i\langle \xi, X_t \rangle}] = \exp(t \psi(\xi))$$

$$\psi(\xi) = i\langle \gamma, \xi \rangle - (1/2)\langle \xi, \Sigma \xi \rangle + \int (e^{i\langle \xi, x \rangle} - 1 - i\langle \xi, x \rangle \mathbb{1}_{\{|x| \leq 1\}}) \nu(dx)$$

Il simbolo di Lévy ψ determina completamente la legge del processo.

4.2 Costruzione di Processi di Lévy su Varietà

Estendere i processi di Lévy a varietà curve è delicato: non possiamo semplicemente "sommare salti" poiché punti su una varietà non possono essere addizionati. La costruzione richiede l'uso del trasporto parallelo e delle mappe esponenziali.

Definizione (Processo di Lévy su M): Un processo X_t su una varietà riemanniana (M, g) è detto di Lévy se:

1. Ha incrementi stazionari: la legge di $\log_{X_t}(X_{t+s})$ dipende solo da s
2. Ha incrementi indipendenti: $\log_{X_{t_1}}(X_{t_2})$ e $\log_{X_{t_2}}(X_{t_3})$ sono indipendenti per $t_1 < t_2 < t_3$

La seconda condizione usa la mappa logaritmica per "trasportare" l'incremento nello spazio tangente dove l'indipendenza ha senso.

Un approccio costruttivo procede via processo di Poisson composto. Dato un processo di Poisson N_t con intensità λ , e una sequenza i.i.d. di "direzioni di salto" $\{V_i\}$ con $V_i \in T_{X_{\tau_i^-}}M$ dove τ_i sono i tempi di salto:

$$X_t = \exp_{X_{\tau_{N_t}^-}} \left(\sum_{i=1}^{N_t} P_{\tau_i \rightarrow t}(V_i) \right)$$

Tra i salti, il processo evolve secondo un moto browniano con drift:

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t \text{ per } t \notin \{\tau_i\}$$

Al tempo di salto τ_i , il processo "salta" da $X_{\tau_i^-}$ a $X_{\tau_i} = \exp_{X_{\tau_i^-}}(V_i)$.

4.3 Il Generatore Infinitesimale e l'Equazione Integro-Differenziale

Il generatore infinitesimale di un processo di Lévy su M ha la forma generale:

$$\mathcal{L}f(x) = \langle b, \nabla f \rangle + (1/2)\text{Tr}(a \text{ Hess } f) + \int_{T_x M} [f(\exp_x(v)) - f(x) - \langle v, \nabla f(x) \rangle \mathbb{1}_{\{|v| \leq 1\}}] \nu_x(dv)$$

Il terzo termine è l'integrale di salto, che cattura i contributi delle discontinuità. La misura di Lévy ν_x su $T_x M$ specifica l'intensità e distribuzione dei salti da x .

Il processo X_t risolve l'equazione integro-differenziale stocastica (EIDS):

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t + \int_{T_{X_{t-}} M} v [N(dt, dv) - \nu_{X_{t-}}(dv) dt]$$

dove $N(dt, dv)$ è la misura aleatoria di Poisson che conta i salti, e il termine compensato $N - \nu dt$ garantisce che il processo sia una martingala locale.

Esempio fondamentale: Il processo di Poisson composto sferico

Sulla sfera S^2 , consideriamo un processo che evolve come moto browniano tra salti casuali. I salti avvengono secondo un processo di Poisson con intensità λ , e ad ogni salto la particella si muove di una distanza geodetica aleatoria $D \sim \rho(\cdot)$ in una direzione uniforme.

Il generatore è:

$$\mathcal{L}f(p) = (1/2)\Delta_g f(p) + \lambda \int_0^\infty \pi \int_{S^1} [f(\exp_p(r\theta)) - f(p)] \rho(r) (dr d\theta) / (2\pi)$$

dove θ parametrizza le direzioni in $T_p S^2$ e ρ è la densità della distanza di salto.

Per salti piccoli (ρ concentrata vicino a 0), il processo approssima una diffusione con coefficiente di diffusione aumentato. Per salti grandi, emerge comportamento qualitativo diverso: la distribuzione stazionaria può essere non uniforme anche su varietà omogenee.

4.4 Subordinazione Temporale e Processi Variance-Gamma

Una tecnica potente per costruire processi di Lévy è la **subordinazione**: dato un processo subordinatore τ_t (processo di Lévy crescente) e un processo di base X_t su M , si definisce:

$$Y_t = X_{\tau_t}$$

Il processo subordinato Y_t "corre" al tempo aleatorio τ_t anziché al tempo fisico t . Questa costruzione permette di generare code pesanti e salti da processi continui.

Il processo Variance-Gamma sferico: Sia B_t un moto browniano sulla sfera S^2 e sia τ_t un subordinatore gamma:

$$\mathbb{P}(\tau_t \in ds) = (1/(\Gamma(t/v) v^{t/v})) s^{t/v - 1} e^{-s/v} ds$$

Il processo VG sferico è definito come:

$$X_t^{\{VG\}} = B_{\tau_t}$$

Il parametro $v > 0$ controlla la varianza del subordinatore: per $v \rightarrow 0$, si ottengono traiettorie con salti frequenti; per $v \rightarrow \infty$, si recupera il moto browniano continuo.

La densità di transizione $p_t^{\{VG\}}(x, y)$ del processo VG può essere espressa come miscela delle densità browniane:

$$p_t^{\{VG\}}(x, y) = \int_0^\infty p_s^{\{BM\}}(x, y) (1/(\Gamma(t/v) v^{t/v})) s^{t/v - 1} e^{-s/v} ds$$

dove $p_s^{\{BM\}}(x, y)$ è il nucleo del calore browniano sulla sfera.

Applicazione cartografica: Il processo VG è utile per modellare fenomeni con "burst" di attività seguiti da periodi di quiete, come eventi sismici lungo faglie tettoniche o variazioni rapide di erosione costiera durante tempeste.

4.5 Processi Stabili su Varietà e Attrattori di Lévy

I processi α -stabili su $\mathbb{R}^n$ sono caratterizzati da code di probabilità pesanti con decadimento polinomiale $\sim |x|^{-(n+\alpha)}$ per $\alpha \in (0, 2)$. Su varietà, la costruzione richiede particolare attenzione alla geometria globale.

Definizione (Processo α -stabile isotropico su S^2): Un processo X_t su S^2 è detto α -stabile isotropico se la sua misura di Lévy in coordinate geodetiche polari centrate in qualsiasi punto ha la forma:

$$\nu_p(dr, d\theta) = (c_\alpha)/(r^{1+\alpha}) dr (d\theta)/(2\pi)$$

per $0 < \alpha < 2$, dove c_α è una costante di normalizzazione e θ parametrizza la direzione del salto.

La proprietà chiave è l'invarianza per rotazioni: la legge del processo è invariante sotto $SO(3)$.

Teorema (Convergenza di somme rinormalizzate): Siano $\{V_i\}$ variabili aleatorie i.i.d. a valori in $T_p S^2$ con code pesanti:

$$\mathbb{P}(|V_i| > r) \sim L(r) r^{-\alpha}$$

dove $L(\cdot)$ è a variazione lenta. Definendo la camminata aleatoria sulla sfera:

$$S_n = \exp_p(\sum_{i=1}^n V_i \text{ con trasporto parallelo})$$

allora opportunamente rinormalizzata:

$$n^{-1/\alpha} S_n \Rightarrow X_1$$

dove X_t è un processo α -stabile su S^2 .

Questo teorema generalizza il teorema del limite centrale per attrattori non gaussiani su varietà curve, con importanti implicazioni per modelli di aggregazione spaziale.

4.6 Processi Jump-Diffusion per Discontinuità Geografiche

Un modello realistico per fenomeni con componente continua e discontinua è il processo jump-diffusion:

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t + dJ_t$$

dove J_t è un processo di salto puro. Questo combina l'evoluzione smooth browniana con transizioni brusche.

Modello di costa con erosione e collasso: Sia $h_t(s)$ l'elevazione della linea di costa al tempo t , parametrizzata dall'arclength s . L'evoluzione è:

$$\partial h_t / \partial t = \kappa \partial^2 h_t / \partial s^2 - v + \sigma \dot{W}_t(s) + \sum_i \Delta_i \delta(s - s_i(t))$$

Componenti:

- $\kappa \partial^2 h_t / \partial s^2$: diffusione (erosione smooth)
- v : drift verso terra (erosione sistematica)
- $\sigma \dot{W}_t(s)$: rumore gaussiano spazio-temporale
- $\sum_i \Delta_i \delta(s - s_i(t))$: collassi improvvisi di scogliere

I collassi avvengono secondo un processo di Poisson spaziale con intensità $\lambda(h_t(s))$ che dipende dall'altezza locale: scogliere più alte hanno maggiore probabilità di collasso.

La distribuzione stazionaria (se esiste) $h_\infty(s)$ bilancia erosione continua e discontinua, portando a profili costieri con geometria frattale.

4.7 La Formula di Lévy-Itô e Integrazione per Processi con Salti

Per un processo generale di Lévy X_t su M con tripla (b, σ, ν) , data $f \in C^2(M)$, la formula di Lévy-Itô generalizza la formula di Itô includendo contributi dei salti:

$$f(X_t) - f(X_0) = \int_0^t \langle b, \nabla f(X_s) \rangle ds + \int_0^t \langle \sigma dB_s, \nabla f(X_s) \rangle + (1/2) \int_0^t \text{Tr}(\sigma \sigma^T \text{Hess } f)(X_s) ds + \int_0^t \int_{T\{X_{s-}\}M} [f(\exp_{X_{s-}}(v)) - f(X_{s-}) - \langle v, \nabla f(X_{s-}) \rangle] \mathbb{1}_{\{|v| \leq 1\}} N(ds, dv) + \int_0^t \int_{T\{X_s\}M} \langle v, \nabla f(X_s) \rangle \mathbb{1}_{\{|v| > 1\}} \nu_{X_s}(dv) ds$$

I primi tre termini sono la parte continua (come in Itô standard), il quarto è la somma dei salti, il quinto è il termine di compensazione che rende il processo una martingala.

Esempio di calcolo: Sulla sfera S^2 con raggio R , consideriamo il processo di Poisson composto con salti di ampiezza fissa d e direzione uniforme. Per la funzione $f(p) = \langle p, n \rangle$ (proiezione su un asse n), abbiamo:

$$\nabla f(p) = (1/R)(n - \langle p, n \rangle p)$$

$$\Delta_g f(p) = -(2/R^2)\langle p, n \rangle$$

Applicando la formula di Lévy-Itô con $b = 0$, $\sigma^2 = \alpha$, salti con intensità λ :

$$f(X_t) - f(X_0) = -\int_0^t (1/R^2)\langle X_s, n \rangle ds + \alpha \int_0^t (1/R)\langle dB_s, n - \langle X_s, n \rangle X_s \rangle + \lambda t \int_{S^1} [f(\exp_{X_0}(d\theta)) - \langle X_0, n \rangle] (d\theta)/(2\pi)$$

dove il termine integrale finale si annulla per simmetria (direzione uniforme).

Capitolo 5: Teoria Ergodica su Varietà e Convergenza di Medie Spaziali

5.1 Ergodicità e Mixing per Processi su Varietà Compatte

La teoria ergodica studia il comportamento a lungo termine di sistemi dinamici. Per processi stocastici su varietà compatte, emergono risultati profondi che collegano medie temporali e medie d'insieme.

Definizione (Misura invariante): Una misura di probabilità μ su M è invariante per il processo $\{X_t\}$ se per ogni insieme misurabile A e ogni $t \geq 0$:

$$\mu(A) = \int_M \mathbb{P}_x(X_t \in A) \mu(dx)$$

Equivalentemente, il semigruppato P_t preserva μ : $\int P_t f d\mu = \int f d\mu$ per ogni $f \in L^1(\mu)$.

Definizione (Ergodicità): Il processo è ergodico rispetto a μ se ogni insieme A invariante (cioè $\mathbb{P}_x(X_t \in A) = \mathbb{1}_A(x)$ per μ -q.o. x e ogni t) ha misura $\mu(A) \in \{0, 1\}$.

Equivalentemente, per ogni $f \in L^2(\mu)$:

$$\lim_{T \rightarrow \infty} (1/T) \int_0^T f(X_t) dt = \int_M f d\mu \text{ q.c. e in } L^2$$

Questa è l'uguaglianza tra media temporale (lato sinistro) e media spaziale (lato destro).

Teorema ergodico di Birkhoff su varietà: Sia X_t un processo ergodico su (M, μ) con μ invariante. Allora per ogni $f \in L^1(\mu)$:

$$\lim_{T \rightarrow \infty} (1/T) \int_0^T f(X_t) dt = \mathbb{E}^\mu[f] \text{ } \mu\text{-q.c.}$$

La convergenza avviene quasi certamente per ogni realizzazione del processo, eccetto un insieme di probabilità zero.

5.2 Tassi di Mixing e Disuguaglianze di Concentrazione

Il mixing quantifica quanto rapidamente un processo "dimentica" le condizioni iniziali. Esistono diverse nozioni di mixing con implicazioni crescenti.

Definizione (Mixing forte): Il processo è mixing forte rispetto a μ se per ogni A, B misurabili:

$$\lim_{t \rightarrow \infty} |\mathbb{P}_\mu(X_0 \in A, X_t \in B) - \mu(A)\mu(B)| = 0$$

Equivalentemente, la correlazione tra osservazioni al tempo 0 e t decade a zero:

$$\lim_{t \rightarrow \infty} |\text{Cov}_\mu(f(X_0), g(X_t))| = 0$$

per ogni $f, g \in L^2(\mu)$.

Teorema (Mixing esponenziale per curvatura positiva): Sia X_t il moto browniano su una varietà compatta (M, g) con curvatura di Ricci $\text{Ric} \geq \kappa g$ per $\kappa > 0$. Allora per ogni $f, g \in L^2(\mu_{\text{unif}})$ con $\int f d\mu = \int g d\mu = 0$:

$$|\text{Cov}_\mu(f(X_0), g(X_t))| \leq C e^{-\lambda t} \|f\|_2 \|g\|_2$$

dove $\lambda > 0$ dipende da κ , dal diametro $\text{diam}(M)$, e dalla dimensione $n = \dim M$.

Il mixing esponenziale è cruciale per applicazioni statistiche: permette di trattare osservazioni distanti nel tempo come approssimativamente indipendenti.

Corollario (Legge dei grandi numeri rafforzata): Sotto le ipotesi del teorema, per $f \in L^2(\mu)$ con $\mathbb{E}[f] = 0$ e decorrelazione esponenziale:

$$\mathbb{P}(|(1/T) \int_0^T f(X_t) dt| > \varepsilon) \leq 2 \exp(-cT\varepsilon^2/\text{Var}(f))$$

per qualche $c > 0$. Questo fornisce controllo esponenziale sulla deviazione dalla media.

5.3 Teorema del Limite Centrale per Additive Functionals

Oltre alla convergenza della media, vogliamo comprendere le fluttuazioni attorno alla media. Il teorema del limite centrale fornisce la risposta.

Teorema (CLT per processi di diffusione ergodici): Sia X_t una diffusione ergodica su M con generatore $\mathcal{L}$ e misura invariante μ . Per $f \in L^2(\mu)$ con $\mathbb{E}^\mu[f] = 0$, se esiste $g \in \text{Dom}(\mathcal{L})$ soluzione dell'equazione di Poisson:

$$\mathcal{L}g = -f$$

con $\int g \, d\mu = 0$, allora:

$$(1/\sqrt{T}) \int_0^T f(X_t) \, dt \Rightarrow \mathcal{N}(0, \sigma^2)$$

$$\text{dove } \sigma^2 = -2 \int_M f g \, d\mu = 2 \int_M g \mathcal{L}g \, d\mu.$$

La varianza asintotica σ^2 può essere calcolata risolvendo l'equazione di Poisson sul generatore. Per il moto browniano su S^2 con f a media nulla:

$$\sigma^2 = -2 \langle f, \Delta_g^{-1} f \rangle_{L^2}$$

dove Δ_g^{-1} è l'operatore inverso del Laplaciano ristretto alle funzioni a media nulla.

Applicazione: Stima dell'elevazione media da campionamento sparso

Supponiamo di osservare l'elevazione topografica $h: S^2 \rightarrow \mathbb{R}$ lungo una traiettoria X_t (ad esempio, un satellite in orbita). Vogliamo stimare l'elevazione media globale:

$$\mu_h = (1/(4\pi R^2)) \int_{S^2} h(p) \, dV_g(p)$$

Lo stimatore naturale è:

$$\hat{\mu}_h(T) = (1/T) \int_0^T h(X_t) \, dt$$

Il teorema ergodico garantisce $\hat{\mu}_h(T) \rightarrow \mu_h$ q.c. Il CLT fornisce l'errore:

$$\sqrt{T} (\hat{\mu}_h(T) - \mu_h) \Rightarrow \mathcal{N}(0, \sigma_h^2)$$

con σ_h^2 determinata dalla struttura di correlazione di h e dalla dinamica di X_t .

Per satellite in orbita circolare con velocità angolare ω , se h ha correlazione esponenziale con scala ℓ , la varianza asintotica è approssimativamente:

$$\sigma_h^2 \approx 2\ell\omega \text{Var}(h)$$

Questo quantifica l'efficienza del campionamento: maggiore è ℓ (correlazione a lungo raggio), maggiore è σ^2 (minor precisione per tempo fissato).

5.4 Principio di Invarianza e Convergenza a Processi Gaussiani

Oltre alla convergenza in distribuzione del valore finale, possiamo studiare la convergenza dell'intero processo rescaled al continuo.

Teorema (Principio di invarianza di Donsker su varietà): Sia $\{X_n\}$ una catena di Markov ergodica su M con misura invariante μ . Per $f: M \rightarrow \mathbb{R}$ con $\mathbb{E}[f] = 0$, definiamo il processo poligonale rescaled:

$$W_T(s) = (1/\sqrt{T}) \sum_{k=1}^{\lfloor Ts \rfloor} f(X_k) \text{ per } s \in [0, 1]$$

Allora:

$$W_T \Rightarrow W$$

dove W è un moto browniano standard con varianza determinata dalla serie:

$$\sigma^2 = \text{Var}(f) + 2 \sum_{k=1}^{\infty} \text{Cov}(f(X_0), f(X_k))$$

La convergenza è nella topologia di Skorokhod su $D[0,1]$, lo spazio delle funzioni càdlàg.

Questo risultato giustifica l'uso di approssimazioni gaussiane per funzionali additivi anche quando le osservazioni singole non sono gaussiane, purché siano sufficientemente decorrelate.

5.5 Teoria Spettrale e Velocità di Convergenza

La velocità di convergenza nell'ergodicità è controllata dallo spectral gap del generatore infinitesimale.

Definizione (Spectral gap): Per una diffusione con generatore $\mathcal{L}$ auto-aggiunto su $L^2(\mu)$, lo spectral gap è:

$$\lambda_1 = \inf\{ \langle f, \mathcal{L}f \rangle : \|f\|_2 = 1, \int f d\mu = 0 \}$$

Lo spectral gap è l'autovalore positivo più piccolo di $-\mathcal{L}$. Controlla il tasso di decadimento esponenziale verso l'equilibrio.

Teorema (Disuguaglianza di Poincaré): Se $\lambda_1 > 0$, allora per ogni $f \in H^1(M)$:

$$\text{Var}_\mu(f) \leq (1/\lambda_1) \int_M \|\nabla f\|^2 d\mu$$

La disuguaglianza di Poincaré quantifica quanto le funzioni devono variare spazialmente per avere varianza positiva. È equivalente alla convergenza esponenziale verso μ con rate λ_1 .

Calcolo esplicito sulla sfera: Per il moto browniano su S^2 di raggio R , gli autovalori del Laplaciano sono $\lambda_\ell = \ell(\ell+1)/R^2$. Lo spectral gap è:

$$\lambda_1 = 2/R^2$$

Questo implica che la distanza in variazione totale tra $p_t(x, \cdot)$ e la misura uniforme decade come:

$$\|p_t(x, \cdot) - \mu_{\text{unif}}\|_{TV} \leq C e^{-t/R^2}$$

Per l'ellissoide terrestre con $R \approx 6371$ km, il tempo caratteristico di mixing è $\sim R^2 \approx 4 \times 10^7$ km², che in termini della velocità tipica di un processo di diffusione atmosferico fornisce scale temporali osservabili.

5.6 Ergodic Theorems per Osservazioni su Griglia Regolare

In pratica, le osservazioni cartografiche sono spesso acquisite su una griglia regolare anziché lungo traiettorie continue. La teoria ergodica si estende a questo contesto discreto.

Setup: Consideriamo una griglia di punti $\{p_i\}_{i \in \mathbb{Z}^2}$ sulla sfera, distribuiti approssimativamente uniformemente con mesh δ . Osserviamo un campo $h: S^2 \rightarrow \mathbb{R}$ nei punti della griglia: $\{h(p_i)\}$.

Se h è realizzazione di un campo ergodico (nel senso che la sua distribuzione è invariante per rotazioni e il processo spaziale associato è mixing), possiamo applicare teoremi ergodici spaziali.

Teorema (Legge dei grandi numeri spaziale): Sia h un campo aleatorio ergodico su S^2 rispetto alle rotazioni. Allora per regioni $A_n \subset S^2$ che "crescono all'infinito" (in senso opportuno):

$$\lim_{n \rightarrow \infty} (1/\text{Area}(A_n)) \int_{A_n} h(p) dV_g(p) = \mathbb{E}[h] \text{ q.c.}$$

La crescita "all'infinito" sulla sfera compatta si interpreta come sequenze di regioni $\{A_n\}$ la cui unione ricopre S^2 in modo sempre più fine e uniforme.

Per campi gaussiani isotropici, si può dimostrare il CLT spaziale:

$$(1/\sqrt{\text{Area}(A_n)}) (\int_{A_n} h(p) dV_g(p) - \mathbb{E}[h] \text{Area}(A_n)) \Rightarrow \mathcal{N}(0, \sigma^2_{\text{spatial}})$$

dove $\sigma^2_{\text{spatial}}$ dipende dalla funzione di correlazione integrata:

$$\sigma^2_{\text{spatial}} = \int_{S^2} C(0, p) dV_g(p)$$

5.7 Applicazione: Validazione di Modelli Climatici Globali

I modelli climatici numerici (GCM - General Circulation Models) producono campi spazio-temporali su S^2 (temperatura, precipitazione, pressione). La teoria ergodica fornisce strumenti per validare questi modelli confrontando:

1. **Medie temporali:** calcolate lungo l'evoluzione del modello
2. **Medie spaziali:** calcolate su ensemble di simulazioni con condizioni iniziali perturbate

Se il modello è ergodico, queste due medie dovrebbero coincidere. Discrepanze indicano:

- Non-ergodicità (presenza di attrattori multipli)
- Transitori lunghi (spectral gap piccolo)
- Bias sistematici nel modello

Procedura diagnostica:

1. Eseguire una simulazione lunga: $T_{\text{lungo}} = 1000$ anni
2. Calcolare la media temporale: $\langle T \rangle_{\text{time}} = (1/T_{\text{lungo}}) \int_0^{T_{\text{lungo}}} T_{\text{global}}(t) dt$
3. Eseguire $N_{\text{ensemble}} = 100$ simulazioni brevi con IC perturbate: $T_{\text{breve}} = 10$ anni
4. Calcolare la media d'ensemble: $\langle T \rangle_{\text{ens}} = (1/N_{\text{ens}}) \sum_{i=1}^{N_{\text{ens}}} [(1/T_{\text{breve}}) \int_0^{T_{\text{breve}}} T_{\text{global}}^{(i)}(t) dt]$
5. Testare: $|\langle T \rangle_{\text{time}} - \langle T \rangle_{\text{ens}}| \ll \sigma/\sqrt{N_{\text{ens}}}$

dove σ è la deviazione standard stimata della media temporale.

La teoria ergodica garantisce che sotto ergodicità, il test dovrebbe passare. Fallimenti suggeriscono problemi nel modello o tempi di simulazione insufficienti.

Capitolo 6: Bootstrap Geometrico e Quantificazione dell'Incertezza

6.1 Il Metodo Bootstrap Classico e le Sue Limitazioni su Varietà

Il bootstrap, introdotto da Efron (1979), è una tecnica di ricampionamento per quantificare l'incertezza statistica senza assumere forme parametriche. Dato un campione $\{x_1, \dots, x_n\}$ e uno stimatore $\theta_n = g(x_1, \dots, x_n)$, il bootstrap genera campioni fittizi $\{x_1, \dots, x_n\}$ ricampionando con reinserimento, e calcola $\theta_n^* = g(x_1^*, \dots, x_n^*)$.

La distribuzione empirica di $\{\theta_n^{*(b)}\}_{b=1}^B$ (su B ricampionamenti) approssima la distribuzione campionaria di θ_n .

Problema su varietà: Il bootstrap classico assume che i dati vivano in uno spazio vettoriale dove ha senso ricampionare. Su una varietà curva M , non possiamo semplicemente "copiare" punti e creare nuovi campioni senza considerare la geometria.

Difficoltà specifiche:

1. **Assenza di struttura lineare:** non possiamo combinare linearmente punti su M
2. **Curvatura:** le medie empiriche standard (baricentri) dipendono dalla metrica
3. **Topologia non banale:** su varietà non semplicemente connesse, il ricampionamento può creare configurazioni inconsistenti

6.2 Bootstrap su Varietà: Approccio via Trasporto nei Tangent Spaces

L'idea chiave è "linearizzare" localmente la varietà mappando i dati in uno spazio tangente, applicare bootstrap euclideo, e riportare i risultati sulla varietà.

Algoritmo (Tangent space bootstrap):

Dato campione $\{p_1, \dots, p_n\} \subset M$ e stimatore $\theta = g(p_1, \dots, p_n) \in \mathbb{R}^d$:

1. Calcolare la media di Fréchet: $\bar{p} = \operatorname{argmin}_{\{p \in M\}} \sum_{i=1}^n d^2(p, p_i)$
2. Mappare i dati nello spazio tangente: $v_i = \log_{\{\bar{p}\}}(p_i) \in T_{\{\bar{p}\}}M$
3. Per $b = 1, \dots, B$: a. Ricampionare con reinserimento: $\{v_1, \dots, v_n\}$ da $\{v_1, \dots, v_n\}$ b. Mappare sulla varietà: $p_i = \exp_{\{\bar{p}\}}(v_i)$ c. Calcolare stimatore bootstrap: $\theta^*(b) = g(p_1, \dots, p_n)$
4. Usare $\{\theta^*(b)\}$ per intervalli di confidenza, test, ecc.

Teorema (Consistenza del tangent space bootstrap): Sotto condizioni di regolarità su g e assumendo che i dati siano concentrati in una regione di curvatura limitata attorno a $\bar{p}$, la distribuzione bootstrap:

$$\mathcal{L}^*(\theta^* - \theta \mid \text{dati}) \rightarrow \mathcal{L}(\theta_n - \theta_0)$$

dove $\rightarrow$ denota convergenza in probabilità (rispetto al ricampionamento bootstrap), $\mathcal{L}^*$ è la distribuzione bootstrap, e $\mathcal{L}$ è la distribuzione campionaria vera.

La convergenza richiede che il raggio di injectivity della mappa esponenziale sia sufficientemente grande rispetto alla dispersione dei dati.

6.3 Smooth Bootstrap e Perturbazioni Gaussiane

Un approccio alternativo, particolarmente utile quando i dati sono già distribuiti secondo un campo aleatorio, è il **smooth bootstrap** che aggiunge perturbazioni gaussiane anziché ricampionare discretamente.

Algoritmo (Gaussian perturbation bootstrap su M):

1. Stimare la covarianza empirica nel tangent space: $\Sigma = (1/n) \sum_{i=1}^n v_i \otimes v_i$ dove $v_i = \log_{\{\bar{p}\}}(p_i)$
2. Per $b = 1, \dots, B$: a. Generare perturbazioni: $\varepsilon_i^*(b) \sim \mathcal{N}(0, h^2 \Sigma)$ in $T_{\{\bar{p}\}}M$ b. Perturbare i dati: $p_i = \exp_{\{\bar{p}\}}(v_i + \varepsilon_i^*(b))$ c. Calcolare $\theta^*(b) = g(p_1, \dots, p_n)$
3. La distribuzione $\{\theta^*(b)\}$ fornisce l'incertezza

Il parametro $h > 0$ è la bandwidth, tipicamente $h \sim n^{-1/(4+\dim M)}$.

Questo approccio è più smooth e può avere migliori proprietà asintotiche del secondo ordine rispetto al bootstrap discreto, specialmente per stimatori non lineari.

6.4 Bootstrap per Campi Aleatori: Block Bootstrap Sferico

Per dati spaziali correlati su S^2 (come campi climatici, topografici, o geofisici), il bootstrap standard è inappropriato perché ignora le correlazioni spaziali. Il **block bootstrap** ricampiona blocchi di regioni contigue.

Algoritmo (Spherical block bootstrap):

Dato campo $h: S^2 \rightarrow \mathbb{R}$ osservato su griglia $\{p_i\}_{i=1}^N$:

1. Dividere S^2 in regioni $B_1, \dots, B_K$ (blocchi) di area approssimativamente uguale
2. Per $b = 1, \dots, B$: a. Ricampionare K blocchi con reinserimento: $B_1, \dots, B_K$ b. Costruire campo bootstrap $h^*\{(b)\}$ assemblando i blocchi ricampionati c. Calcolare statistica di interesse: $\theta^*\{(b)\} = g(h^*\{(b)\})$
3. Usare $\{\theta^*\{(b)\}\}$ per inferenza

Scelta della dimensione dei blocchi: I blocchi devono essere:

- Abbastanza grandi da catturare la correlazione spaziale locale
- Abbastanza piccoli da avere molti blocchi (per variabilità bootstrap)

Un criterio empirico: scegliere dimensione blocco $\sim \ell_{\text{corr}}$ dove ℓ_{corr} è la scala di correlazione del campo (distanza oltre cui la correlazione diventa trascurabile).

Teorema (Consistenza del block bootstrap sferico): Sia h un campo aleatorio stazionario su S^2 con mixing $\alpha(r) \rightarrow 0$ per $r \rightarrow \infty$. Se la dimensione dei blocchi soddisfa:

$$\ell_{\text{block}} \rightarrow \infty \text{ e } \ell_{\text{block}}/R_{\text{sphere}} \rightarrow 0 \text{ quando } n \rightarrow \infty$$

allora la distribuzione bootstrap converge:

$$\sup_x |\mathbb{P}^*(\theta^* \leq x) - \mathbb{P}(\theta_n \leq x)| \rightarrow 0$$

La convergenza è in probabilità rispetto ai dati originali.

6.5 Propagazione dell'Incertezza attraverso Proiezioni Cartografiche

Le proiezioni cartografiche $f: S^2 \rightarrow \mathbb{R}^2$ sono trasformazioni non lineari che distorcono aree, angoli, o distanze. Vogliamo quantificare come l'incertezza nei dati sferici si propaga attraverso la proiezione.

Setup: Supponiamo di avere:

- Dati sulla sfera: $\{p_i\} \subset S^2$ con incertezza (ad es., da GPS con errore di misura)
- Proiezione cartografica: $f: S^2 \rightarrow \mathbb{R}^2$
- Quantità di interesse nella mappa piatta: $\theta = g(f(p_1), \dots, f(p_n))$

Metodo 1: Bootstrap parametrico con modello di errore

1. Modellare l'errore di misura: $p_i^{\text{obs}} = \exp\{p_i^{\text{true}}\}(\varepsilon_i)$ dove $\varepsilon_i \sim \mathcal{N}(0, \Sigma_{\text{GPS}})$
2. Stimare posizioni vere: $\hat{p}_i$ via maximum likelihood o filtering
3. Bootstrap: a. Per $b = 1, \dots, B$:
 - Simulare errori: $\varepsilon_i^* \sim \mathcal{N}(0, \Sigma_{\text{GPS}})$
 - Generare osservazioni bootstrap: $p_i = \exp\{\hat{p}_i\}(\varepsilon_i)$
 - Proiettare: $q_i = f(p_i)$
 - Calcolare: $\theta^*\{(b)\} = g(q_1, \dots, q_n)$
4. La distribuzione $\{\theta^*\{(b)\}\}$ quantifica l'incertezza propagata

Metodo 2: Delta method geometrico

Per piccole incertezze, possiamo linearizzare. La mappa di proiezione $f: S^2 \rightarrow \mathbb{R}^2$ ha differenziale:

$$Df_p: T_p S^2 \rightarrow \mathbb{R}^2$$

Se $p \sim \mathcal{N}_M(\mu, \Sigma)$ (distribuzione gaussiana sulla sfera), allora per piccole Σ :

$$f(p) \approx \mathcal{N}(f(\mu), Df_\mu \Sigma Df_\mu^T)$$

La covarianza trasformata $Df_\mu \Sigma Df_\mu^T$ cattura la distorsione della proiezione.

Esempio: Proiezione di Mercatore

La proiezione di Mercatore $(\text{lat}, \text{lon}) \mapsto (x, y)$ è:

$$x = R \text{ lon} \quad y = R \log(\tan(\pi/4 + \text{lat}/2))$$

Il differenziale in un punto $(\text{lat}_0, \text{lon}_0)$ è:

$$Df = R \begin{bmatrix} 1, & 0 \\ 0, & \sec(\text{lat}_0) \end{bmatrix}$$

Una perturbazione circolare sulla sfera ($\Sigma = \sigma^2 I$) diventa un'ellisse allungata verticalmente sulla mappa, con fattore di allungamento $\sec(\text{lat}_0)$ che diverge ai poli.

Per $\text{lat}_0 = 60^\circ$, $\sec(60^\circ) = 2$: un cerchio di raggio σ sulla sfera diventa un'ellisse con assi σ (orizzontale) e 2σ (verticale) sulla mappa.

6.6 Bootstrap Wild per Eteroschedasticità Spaziale

Quando la varianza del campo aleatorio varia spazialmente (eteroschedasticità), il bootstrap standard può sottostimare o sovrastimare l'incertezza in diverse regioni.

Modello: Osserviamo $h(p_i) = \mu(p_i) + \sigma(p_i)\varepsilon_i$ dove $\sigma(\cdot)$ varia spazialmente.

Wild bootstrap algoritmo:

1. Stimare la media: $\hat{\mu}(p_i)$ via smoothing
2. Calcolare residui: $r_i = h(p_i) - \hat{\mu}(p_i)$
3. Stimare varianza locale: $\hat{\sigma}^2(p_i)$ via windowing o kernel smoothing
4. Standardizzare residui: $\tilde{r}_i = r_i / \hat{\sigma}(p_i)$
5. Per $b = 1, \dots, B$: a. Generare moltiplicatori: $w_i \sim \text{Rademacher}(\pm 1 \text{ con prob } 1/2)$ b. Creare residui bootstrap: $r_{_i} = w_{_i} \cdot \tilde{r}_{_i} \cdot \hat{\sigma}(p_{_i})$ c. Ricostruire campo: $h_{_i} = \hat{\mu}(p_{_i}) + r_{_i}$ d. Calcolare statistica: $\theta^{\{b\}} = g(h^*)$
6. Usare $\{\theta^{\{b\}}\}$ per intervalli di confidenza

Il wild bootstrap preserva l'eteroschedasticità stimata mentre permette di ricampionare la variabilità aleatoria.

6.7 Applicazione: Intervalli di Confidenza per Volumi di Calotte Polari

Problema: Stimare il volume di ghiaccio nelle calotte polari (Groenlandia, Antartide) e quantificare l'incertezza, dato:

- Osservazioni altimetriche satellitari (con errore di misura)
- Modello di elevazione del bedrock (con incertezza)
- Copertura spaziale incompleta

Procedura:

1. **Dati:** Elevazioni superficiali $\{h_{\text{surf}}(p_i)\}$ e bedrock $\{h_{\text{bed}}(p_i)\}$ su griglia
2. **Calcolo diretto:** Volume stimato $\hat{V} = \sum_i \text{Area}(\text{cell}_i) \cdot [h_{\text{surf}}(p_i) - h_{\text{bed}}(p_i)]$
3. **Bootstrap spaziale:** a. Modellare entrambi i campi come realizzazioni di processi gaussiani con covarianza stimata b. Per $b = 1, \dots, B$:
 - Simulare realizzazione bootstrap del campo superficie: h^*_{surf} via conditional simulation
 - Simulare realizzazione bootstrap del bedrock: h^*_{bed}
 - Calcolare volume bootstrap: $V^*\{b\} = \sum_i \text{Area}(\text{cell}_i) \cdot [h^*_{\text{surf}}(p_i) - h^*_{\text{bed}}(p_i)]$ c. Intervallo di confidenza al 95%: $[Q\{0.025\}(V^*), Q\{0.975\}(V^*)]$
4. **Correzione per bias di curvatura:** La Terra non è piatta! Correggere gli elementi di area:

$$\text{Area}_{\text{corretta}}(\text{cell}_i) = R^2 \sin(\theta_i) \Delta\theta \Delta\phi$$

dove (θ_i, ϕ_i) sono coordinate sferiche del centro della cella.

Risultati tipici: Per la calotta groenlandese, $\hat{V} \approx 2.85 \times 10^6 \text{ km}^3$ con intervallo bootstrap $[2.82, 2.88] \times 10^6 \text{ km}^3$, riflettendo incertezze nella misura e nella copertura spaziale.

Capitolo 7: Campi Markoviani su Grafi Planari

7.1 Dai Campi Continui alle Rappresentazioni Discrete

Nella pratica cartografica, i dati sono inevitabilmente discretizzati su griglie, mesh triangolari, o strutture di grafi irregolari. Vogliamo estendere la teoria dei campi aleatori continui a questi setting discreti, preservando le proprietà probabilistiche essenziali.

Transizione concettuale:

- Continuo: Campo gaussiano $h: S^2 \rightarrow \mathbb{R}$ con covarianza $C(p, q)$
- Discreto: Vettore aleatorio $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ su vertici di un grafo $G = (V, E)$

La chiave è che la struttura di dipendenza spaziale, codificata dalla covarianza nel continuo, si traduce nella struttura di adiacenza del grafo nel discreto.

7.2 Campi Aleatori di Markov (MRF) su Grafi

Definizione (Proprietà di Markov su grafo): Un campo aleatorio h su vertici V di un grafo G ha la proprietà di Markov se per ogni vertice v :

$$\mathbb{P}(h_v \mid h_{V \setminus \{v\}}) = \mathbb{P}(h_v \mid h_{N(v)})$$

dove $N(v) = \{u \in V : (u,v) \in E\}$ è il vicinato di v . Localmente, h_v dipende solo dai vicini immediati.

Teorema di Hammersley-Clifford: Per distribuzioni positive, la proprietà di Markov è equivalente alla fattorizzazione di Gibbs:

$$\mathbb{P}(h) = (1/Z) \exp(-\sum_{C \in \mathcal{C}} \psi_C(h_C))$$

dove:

- $\mathcal{C}$ è l'insieme delle clique massimali di G
- $\psi_C: \mathbb{R}^{|C|} \rightarrow \mathbb{R}$ sono potenziali di energia
- $Z = \sum_h \exp(-\sum_C \psi_C(h_C))$ è la funzione di partizione

Questa rappresentazione è fondamentale per la modellazione e l'inferenza.

7.3 GMRF: Gaussian Markov Random Fields

Per campi gaussiani, la proprietà di Markov si traduce in una struttura di precisione sparsa.

Definizione (GMRF): Un campo gaussiano $h \sim \mathcal{N}(\mu, \Sigma)$ su un grafo G è un GMRF se la matrice di precisione $Q = \Sigma^{-1}$ riflette la struttura del grafo:

$$Q_{\{ij\}} \neq 0 \Leftrightarrow i = j \text{ oppure } (i,j) \in E$$

Gli elementi non-diagonali di Q sono non-zero solo per vertici adiacenti.

Densità: La distribuzione ha forma chiusa:

$$\begin{aligned} \mathbb{P}(h) &= (1/\sqrt{(2\pi)^n \det(\Sigma)}) \exp(-(1/2)(h - \mu)^T Q (h - \mu)) \\ &= (1/\sqrt{(2\pi)^n}) \sqrt{\det(Q)} \exp(-(1/2) \sum_{\{i,j\} \in E} Q_{\{ij\}}(h_i - \mu_i)(h_j - \mu_j) - (1/2) \sum_i Q_{\{ii\}}(h_i - \mu_i)^2) \end{aligned}$$

La forma quadratica coinvolge solo coppie di vertici adiacenti.

Relazione con campi continui: Un GMRF su una griglia regolare approssima un campo gaussiano continuo. Per una griglia con spacing δ sulla sfera, la matrice di precisione approssima un operatore differenziale:

$$Q \approx -\alpha \Delta_{\text{discrete}} + \beta I$$

dove Δ_{discrete} è il Laplaciano discreto sul grafo e α, β sono parametri.

7.4 Modelli Intrinseci e Prior Impropri

Per modellare campi smooth senza specificare una media fissa, si usano **prior intrinseci** che penalizzano solo le variazioni, non i livelli assoluti.

Il modello CAR intrinseco: Per un grafo con insieme di vertici V e bordi E , il prior CAR (Conditional Autoregressive) intrinseco è:

$$\mathbb{P}(h) \propto \exp(-(\kappa/2) \sum_{\{i,j\} \in E} (h_i - h_j)^2)$$

Questa è una distribuzione impropria (non integra a 1) perché è invariante per traslazioni: h e $h + c\mathbb{1}$ hanno la stessa densità.

La matrice di precisione è:

$$Q_{\{ij\}} = \kappa \cdot (-1 \text{ se } i \sim j, d_i \text{ se } i = j, 0 \text{ altrimenti})$$

dove $d_i = |N(i)|$ è il grado del vertice i .

Interpretazione: Il prior favorisce configurazioni dove vertici adiacenti hanno valori simili. Il parametro $\kappa > 0$ controlla l'intensità della penalizzazione: κ grande $\rightarrow$ campo molto smooth.

Rank deficiency: Q ha rank $n - k$ dove k è il numero di componenti connesse. Per grafi connessi, $\text{rank}(Q) = n - 1$. Il kernel è lo spazio delle funzioni costanti.

Per rendere il prior proprio (necessario per alcune procedure bayesiane), si aggiunge un vincolo di somma-a-zero: $\sum_i h_i = 0$, o una piccola regolarizzazione $Q \rightarrow Q + \epsilon I$.

7.5 Campi SPDE Approssimati: L'Approccio INLA

Un breakthrough per l'inferenza efficiente con campi gaussiani su grafi planari è l'approccio **SPDE (Stochastic Partial Differential Equation)** di Lindgren-Rue-Lindström.

Idea chiave: Rappresentare un campo gaussiano continuo come soluzione di una SPDE, poi discretizzarla via elementi finiti.

Consideriamo l'equazione:

$$(\kappa^2 - \Delta_g)^{\alpha/2} h = \tau W$$

dove:

- W è white noise gaussiano sulla sfera
- $\kappa > 0$ è un parametro di scala
- α determina la smoothness
- $\tau > 0$ controlla la varianza marginale

Soluzione: Per $\alpha = 2$, la soluzione h è un campo gaussiano con covarianza di Matérn:

$$C(p, q) = (\sigma^2 / (\Gamma(v) 2^{v-1})) (\kappa d(p, q))^v K_v(\kappa d(p, q))$$

dove $v = \alpha - d/2$ ($d = 2$ per la sfera), K_v è la funzione di Bessel modificata.

Discretizzazione FEM: Su una triangolazione $\{T_k\}$ della sfera:

1. Rappresentare h mediante funzioni di base a cappello: $h(p) \approx \sum_i h_i \varphi_i(p)$
2. L'equazione SPDE si traduce in un sistema algebrico:

$$(\kappa^2 C + G) h = \tau \tilde{W}$$

dove:

- C è la matrice di massa: $C_{ij} = \int \varphi_i \varphi_j dV_g$
 - G è la matrice di rigidità: $G_{ij} = \int \langle \nabla \varphi_i, \nabla \varphi_j \rangle dV_g$
 - $\tilde{W} \sim \mathcal{N}(0, C)$
3. La matrice di precisione del GMRF risultante è:

$$Q = \tau^{-2} (\kappa^2 C + G) C^{-1} (\kappa^2 C + G)$$

Questa matrice è sparsa (gli elementi non-zero corrispondono a vertici del grafo FEM connessi), permettendo calcoli efficienti anche per $n \sim 10^5 - 10^6$ vertici.

Vantaggi:

- Costo computazionale $O(n^{3/2})$ per grafi planari (vs $O(n^3)$ per metodi diretti)
- Interpretazione probabilistica chiara (campo di Matérn)
- Flessibilità: α, κ, τ possono variare spazialmente

7.6 Condizionamento e Kriging Discreto

Dato un GMRF $h \sim \mathcal{N}(0, Q^{-1})$ e osservazioni rumorose $y = Ah + \varepsilon$ con $\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)$, dove A è una matrice di osservazione, la distribuzione posteriore è:

$$h | y \sim \mathcal{N}(\mu_{\text{post}}, Q_{\text{post}}^{-1})$$

con:

$$Q_{\text{post}} = Q + A^T \Sigma_\varepsilon^{-1} A \quad \mu_{\text{post}} = Q_{\text{post}}^{-1} A^T \Sigma_\varepsilon^{-1} y$$

Osservazione cruciale: Q_{post} è ancora sparsa! Se A è locale (ogni osservazione dipende da pochi vertici), allora $A^T \Sigma_\varepsilon^{-1} A$ aggiunge solo pochi elementi non-zero.

Questo permette **kriging computazionalmente efficiente** su grandi griglie:

1. Fattorizzare $Q_{\text{post}} = LL^T$ (Cholesky sparsa): $O(n^{3/2})$
2. Risolvere $Q_{\text{post}} \mu_{\text{post}} = A^T \Sigma_\varepsilon^{-1} y$ y via back-substitution: $O(n^{3/2})$
3. Simulare da posteriore: $h^* = \mu_{\text{post}} + L^{-T} z$ con $z \sim \mathcal{N}(0, I)$: $O(n)$

7.7 Stima dei Parametri: Maximum Likelihood Approssimato

Per modelli GMRF, la likelihood è:

$$\ell(\theta) = -(1/2) \log \det(2\pi Q(\theta)^{-1}) - (1/2) h^T Q(\theta) h$$

dove $\theta = (\kappa, \tau, \dots)$ sono parametri del modello.

Il termine $\log \det(Q^{-1})$ è costoso da calcolare per grandi n . **Approssimazioni:**

1. **Composite likelihood:** Usare solo sotto-modelli locali
2. **Laplace approximation:** Per modelli non-gaussiani
3. **REML (Restricted Maximum Likelihood):** Elimina effetti di media

INLA (Integrated Nested Laplace Approximations): Un framework computazionale che combina:

- SPDE per rappresentare campi continui
- Approssimazioni di Laplace per marginali posteriori
- Metodi sparsi per algebra lineare

Permette inferenza bayesiana completa (posteriori per h e θ) in tempi dell'ordine di minuti per modelli con $n \sim 10^4 - 10^5$ vertici.

7.8 Applicazione: Mappatura di Malattie con Dati Aggregati su Regioni

Setup: Abbiamo conteggi di casi di malattia Y_i in regioni geografiche R_i (ad es., province), con popolazioni N_i . Vogliamo stimare il rischio spaziale continuo $p: S^2 \rightarrow \mathbb{R}_+$.

Modello gerarchico:

1. **Livello osservazione:** $Y_i | \lambda_i \sim \text{Poisson}(N_i \lambda_i)$
2. **Livello spaziale:** $\log(\lambda_i) = \beta_0 + \beta_1 X_i + h_i$
 - β_0 : baseline log-risk
 - $\beta_1 X_i$: effetti di covariate (età, densità, ...)
 - h_i : effetto spaziale aleatorio
3. **Prior spaziale:** $h \sim \text{GMRF}$ con struttura CAR sul grafo di adiacenza delle regioni:

$$Q_{\{ij\}} = (-\kappa \text{ se regioni } i, j \text{ adiacenti, } \kappa \cdot d_i \text{ se } i = j)$$
4. **Prior sui parametri:** $\beta \sim \mathcal{N}(0, \sigma^2 \beta I)$, $\kappa \sim \text{Gamma}(a, b)$

Inferenza via INLA:

1. Discretizzare S^2 con mesh fine
2. Aggregare il campo h dalla mesh alle regioni: $h_i = (1/|R_i|) \int_{R_i} h(p) dp$
3. Calcolare approssimazioni di Laplace per marginali posteriori $\mathbb{P}(h_i | Y)$, $\mathbb{P}(\kappa | Y)$

4. Predire il rischio in nuove location: $\mathbb{P}(\lambda_{\text{new}} | Y)$

Output:

- Mappe di rischio posteriore: $\mathbb{E}[\lambda_i | Y]$ con intervalli di credibilità
- Identificazione di hotspot: regioni con $\mathbb{P}(\lambda_i > \text{threshold} | Y) > 0.95$
- Quantificazione dell'incertezza spaziale

Esempio reale: Mappatura dell'incidenza di malaria in Africa sub-sahariana. Con $\sim 10^4$ regioni e $\sim 10^5$ vertici nella mesh, INLA produce stime posteriori in ~ 10 minuti su workstation standard, permettendo analisi iterativa e model selection.

Capitolo 8: Teoria dei Grandi Deviazioni e Eventi Estremi su Sfere

8.1 Principi di Grandi Deviazioni per Campi Gaussiani

La teoria dei grandi deviazioni quantifica le probabilità di eventi rari per campi aleatori, essenziale per valutare rischi di soglie critiche (inondazioni, onde anomale, picchi topografici estremi).

Setup: Sia h_ε un campo gaussiano centrato su S^2 dipendente da un parametro $\varepsilon \rightarrow 0$ (ad es., intensità del rumore). Vogliamo studiare:

$$\mathbb{P}(\sup_{p \in S^2} h_\varepsilon(p) \geq a)$$

per a grande, quando $\varepsilon \rightarrow 0$.

Principio LDP (Large Deviation Principle): Se esiste una funzione $I: C(S^2) \rightarrow [0, \infty]$ (rate function) tale che per ogni insieme A :

$$-\inf_{f \in \text{Int}(A)} I(f) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(h_\varepsilon \in A) \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon^2 \log \mathbb{P}(h_\varepsilon \in A) \leq -\inf_{f \in \text{Cl}(A)} I(f)$$

allora $\{h_\varepsilon\}$ soddisfa un LDP con rate function I e velocità ε^{-2} .

Forma esplicita per GMRF: Per $h_\varepsilon = \varepsilon W$ con $W \sim \mathcal{N}(0, Q^{-1})$:

$$I(f) = (1/2) f^T Q f$$

Gli eventi con $I(f)$ piccola sono esponenzialmente più probabili.

8.2 Distribuzioni di Estremi per Campi Lisci

Per campi smooth, la distribuzione del massimo $\sup_{p \in M} h(p)$ è determinata da risultati classici nella teoria degli estremi gaussiani.

Teorema (Distribuzione asintotica del massimo): Sia h un campo gaussiano isotropico su S^2 con varianza σ^2 e funzione di correlazione $\rho(\theta)$. Sotto condizioni di regolarità (decadimento sufficiente di ρ), per u grande:

$$\mathbb{P}(\sup_{p \in S^2} h(p) \leq u) \approx \exp(-4\pi R^2 c(\rho) \Phi(u/\sigma))$$

dove:

- $\Phi(x) = \mathbb{P}(Z > x)$ per $Z \sim \mathcal{N}(0,1)$ è la coda gaussiana
- $c(\rho)$ è la **costante di Pickands** dipendente dalla curvatura locale della correlazione

Per correlazione gaussiana $\rho(\theta) = \exp(-\theta^2/(2\ell^2))$, la costante è esplicitamente:

$$c(\rho) = (1/(2\pi)) \sqrt{\rho''(0)} = 1/(2\pi\ell^2)$$

Interpretazione: La probabilità che il massimo ecceda una soglia alta u decresce esponenzialmente con tasso controllato da:

1. L'area della sfera ($4\pi R^2$)
2. La scala di correlazione (tramite $c(\rho)$)
3. La coda gaussiana $\Phi(u/\sigma) \approx (\sigma/u) \phi(u/\sigma)$ con ϕ densità standard

8.3 Crossing Probabilities e Rice Formula

Un problema fondamentale: qual è la probabilità che un campo aleatorio attraversi un livello u in almeno un punto?

Rice formula per campi differenziabili: Sia h un campo gaussiano C^1 su S^2 con derivate gaussiane congiunte. Il numero atteso di upcrossing del livello u (punti dove $h(p) = u$ e $\nabla h(p)$ punta verso l'alto) è:

$$\mathbb{E}[N_u] = \int_{S^2} \mathbb{E}[|\nabla h(p)|^+ | h(p) = u] \cdot p_h(u) dV_g(p)$$

dove:

- $p_h(u)$ è la densità marginale di $h(p)$
- $|\nabla h|^+ = \max(\langle \nabla h, n \rangle, 0)$ per qualche campo di direzioni n

Per campi isotropici gaussiani con $h(p) \sim \mathcal{N}(0, \sigma^2)$ e ∇h avente distribuzione indipendente da h :

$$\mathbb{E}[N_u] = (4\pi R^2)/(2\pi) \sqrt{\lambda_2} \exp(-u^2/(2\sigma^2))$$

dove $\lambda_2 = \mathbb{E}[\|\nabla h\|^2]$ è correlata alla curvatura della covarianza.

Applicazione: Conteggio di picchi topografici: Per un modello di elevazione digitale (DEM) modellato come campo gaussiano, il numero atteso di picchi locali (massimi isolati) sopra quota h_{crit} è approssimativamente $\mathbb{E}[N_{h_{\text{crit}}}] / 2$.

8.4 Teoria dei Valori Estremi per Osservazioni su Griglia

Nella pratica, osserviamo h su un insieme finito di punti $\{p_1, \dots, p_n\}$. La teoria GEV (Generalized Extreme Value) caratterizza la distribuzione dei massimi campionari.

Teorema di Fisher-Tippett: Sia $M_n = \max\{h(p_1), \dots, h(p_n)\}$ il massimo su n punti. Se esistono sequenze $a_n > 0$, b_n tali che:

$$(M_n - b_n)/a_n \Rightarrow G$$

per qualche distribuzione limite G non degenera, allora G è una delle tre famiglie GEV:

1. **Gumbel:** $G(x) = \exp(-\exp(-(x-\mu)/\sigma))$ per $x \in \mathbb{R}$
2. **Fréchet:** $G(x) = \exp(-(x/\sigma)^{-\alpha})$ per $x > 0$, $\alpha > 0$
3. **Weibull:** $G(x) = \exp(-((x-\mu)/\sigma)^\alpha)$ per $x < \mu$, $\alpha > 0$

Per campi gaussiani con correlazione decadente sufficientemente rapida, il limite è **Gumbel**.

Parametrizzazione unificata: La famiglia GEV unificata è:

$$G(x; \mu, \sigma, \xi) = \exp(-(1 + \xi(x-\mu)/\sigma)^{-1/\xi})$$

per $1 + \xi(x-\mu)/\sigma > 0$, dove:

- μ : parametro di posizione
- $\sigma > 0$: parametro di scala
- ξ : parametro di forma ($\xi = 0 \rightarrow$ Gumbel, $\xi > 0 \rightarrow$ Fréchet, $\xi < 0 \rightarrow$ Weibull)

Stima: Dati massimi campionari $\{M_1, \dots, M_k\}$ (ad es., massimi annuali di temperatura), stimare (μ, σ, ξ) via maximum likelihood o method of moments.

8.5 Modelli POT (Peaks Over Threshold)

Un approccio alternativo più efficiente per dati limitati: modellare tutti i valori sopra una soglia alta u_0 , non solo i massimi per blocchi.

Teorema di Pickands-Balkema-de Haan: Per una soglia u_0 sufficientemente alta, la distribuzione condizionale degli eccessi oltre u_0 converge a una Generalized Pareto Distribution (GPD):

$$F_{\{u_0\}}(y) = \mathbb{P}(h - u_0 \leq y \mid h > u_0) \rightarrow G_{\{GP\}}(y; \sigma_u, \xi)$$

dove:

$$G_{\{GP\}}(y; \sigma, \xi) = 1 - (1 + \xi y/\sigma)^{-1/\xi}$$

per $y \geq 0$ quando $\xi \geq 0$, e $0 \leq y \leq -\sigma/\xi$ quando $\xi < 0$.

Procedura di stima:

1. Scegliere soglia u_0 (compromesso: troppo bassa $\rightarrow$ bias, troppo alta $\rightarrow$ pochi dati)
2. Estrarre eccessi: $y_i = h(p_i) - u_0$ per tutti $i \in \{i : h(p_i) > u_0\}$
3. Fittare GPD ai $\{y_i\}$ via ML: (σ, ξ)
4. Predire quantili estremi: per probabilità di eccedenza piccola p ,

$$x_p = u_0 + (\sigma/\xi)[(N_u p/N)^{-\xi} - 1]$$

dove $N_u = \#\{i : h(p_i) > u_0\}$ e N è il totale delle osservazioni.

8.6 Return Levels e Return Periods su Sfere

Un **return level** è il valore che ci aspettiamo venga superato in media una volta ogni T unità temporali (o spaziali).

Definizione: Il return level con periodo T è il quantile z_T tale che:

$$\mathbb{P}(M_T > z_T) = 1/T$$

dove M_T è il massimo su T osservazioni (o su area equivalente).

Per modello GEV con parametri (μ, σ, ξ) , il return level è:

$$z_T = \mu - (\sigma/\xi)[1 - (-\log(1 - 1/T))^{\xi}]$$

Per $\xi \rightarrow 0$ (Gumbel):

$$z_T = \mu - \sigma \log(-\log(1 - 1/T)) \approx \mu + \sigma \log(T)$$

Interpretazione cartografica: Per precipitazioni estreme, "l'evento centennale" ($T = 100$ anni) corrisponde al quantile z_{100} . Mappe di z_T per $T = 10, 50, 100$ anni visualizzano il rischio spaziale di eventi rari.

8.7 Campi Max-Stable e Modelli Spaziali per Estremi

Per modellare estremi spazialmente correlati, si usano **campi max-stable**: l'analogo spaziale delle distribuzioni GEV.

Definizione: Un campo $h: S^2 \rightarrow \mathbb{R}$ è max-stable se per ogni $n \in \mathbb{N}$ esistono funzioni $a_n(p) > 0$, $b_n(p)$ tali che:

$$\max\{h_1(p), \dots, h_n(p)\} \stackrel{d}{=} a_n(p) h(p) + b_n(p)$$

dove $h_1, \dots, h_n$ sono copie i.i.d. di h e $\stackrel{d}{=}$ denota uguaglianza in distribuzione.

Rappresentazione di de Haan: Ogni campo max-stable può essere rappresentato come:

$$h(p) = \max_{i \geq 1} [U_i \cdot f_i(p)]$$

dove $\{U_i\}$ è un processo di Poisson con intensità $u^{-2} du$ su $(0, \infty)$, e $\{f_i\}$ sono realizzazioni i.i.d. di un processo ausiliario.

Modello di Smith (Gaussian extreme value): Una costruzione esplicita usa funzioni di tempesta gaussiane:

$$h(p) = \max_{i \geq 1} [U_i \cdot \exp(-d^2(p, C_i)/(2\sigma^2))]$$

dove C_i sono centri aleatori uniformi su S^2 . Questo modella estremi come risultato di eventi estremi locali (tempeste) che si sovrappongono.

Inferenza: La likelihood per campi max-stable su griglia finita ha forma complicata (somme su tutti i sottoinsiemi). Metodi pratici:

1. **Composite likelihood:** usare solo coppie o triple di siti
2. **Censored likelihood:** condizionare su eventi estremi osservati
3. **Simulazione:** MCMC con data augmentation per $\{U_i, C_i, f_i\}$

8.8 Applicazione: Mappe di Rischio per Eventi Estremi Climatici

Problema: Costruire mappe di probabilità di superamento di soglie critiche per onde di calore, alluvioni, o siccità.

Dati: Serie temporali di temperatura massima giornaliera $T(p, t)$ su griglia spaziale $\{p_i\}$ e tempo $t = 1, \dots, T$.

Procedura:

1. **Estrazione di blocchi massimi:** Per ogni sito i e anno j : $M_{ij} = \max_{\{t \in \text{anno}_j\}} T(p_i, t)$
2. **Fitting spaziale GEV:** Assumere parametri GEV smooth in spazio: $\mu(p) = \beta_0 + \beta_1 \text{lat}(p) + \beta_2 \text{lon}(p) + h(p)$ $\sigma(p) = \exp(\gamma_0 + \gamma_1 \text{lat}(p))$ $\xi = \xi_0$ (costante)

dove $h(p)$ è un effetto spaziale aleatorio GMRF
3. **Stima bayesiana:** Via INLA o MCMC, ottenere posteriori per tutti i parametri
4. **Predizione di return levels:** Per ogni p nella griglia fine: $z_T(p) \mid \text{dati} \sim \text{posteriore}$
 $\mathbb{P}(z_T(p) > \text{threshold} \mid \text{dati})$
5. **Visualizzazione:** Mappe di:
 - $\mathbb{E}[z_{100}(p) \mid \text{dati}]$: valore atteso dell'evento centennale
 - $\mathbb{P}(z_{100}(p) > 45^\circ\text{C} \mid \text{dati})$: probabilità di superare soglia critica
 - Intervalli di credibilità: $[Q_{0.05}(z_{100}), Q_{0.95}(z_{100})]$

Output: Mappe ad alta risoluzione che identificano regioni con alto rischio di eventi estremi, includendo quantificazione dell'incertezza. Utili per pianificazione infrastrutturale, assicurazioni, e politiche di adattamento climatico.

Capitolo 9: Teoria della Percolazione e Connettività su Grafi Planari

9.1 Percolazione su Reticoli e Grafi Geografici

La teoria della percolazione studia la formazione di cluster connessi in sistemi aleatori, con applicazioni a reti di trasporto, diffusione di epidemie, e connettività geografica.

Setup classico (bond percolation): Su un grafo $G = (V, E)$, ogni arco $e \in E$ è aperto con probabilità p (indipendentemente), chiuso con probabilità $1 - p$. Studiamo:

- La dimensione del cluster contenente un vertice
- Esistenza di cluster infiniti (su grafi infiniti)
- Probabilità di connessione tra punti distanti

Probabilità critica: Esiste $p_c \in (0, 1)$ tale che:

- $p < p_c$: solo cluster finiti q.c.
- $p > p_c$: esiste un cluster infinito con probabilità 1

Per reticolo quadrato bidimensionale: $p_c = 1/2$ esattamente (risultato profondo di Kesten, 1980).

Transizione di fase: A $p = p_c$ il sistema subisce una transizione di fase critica:

- Dimensione del cluster più grande $\sim n^{\beta_p}$
- Lunghezza di correlazione $\sim |p - p_c|^{-\nu}$

con esponenti critici universali che dipendono solo dalla dimensione, non dai dettagli microscopici del sistema.

9.2 Percolazione Continua su Sfere

Su una varietà continua come S^2 , la percolazione continua considera deposizioni aleatorie di regioni geometriche.

Modello Boolean: Posizionare centri C_i secondo un processo di Poisson con intensità λ sulla sfera, e associare a ogni centro una regione aleatoria R_i (ad es., disco di raggio r aleatorio). Studiare la regione coperta:

$$\mathcal{C} = \bigcup_{i \geq 1} R_i$$

e le sue proprietà di connettività.

Soglia di percolazione: Per dischi di raggio fisso r , esiste $\lambda_c(r)$ tale che:

- $\lambda < \lambda_c$: $\mathcal{C}$ consiste in componenti connesse finite q.c.
- $\lambda > \lambda_c$: $\mathcal{C}$ contiene una componente che avvolge la sfera

Su S^2 con raggio R , se i dischi hanno raggio r , la soglia critica è approssimativamente:

$$\lambda_c(r) \approx 1/(\pi r^2) \cdot \log(4\pi R^2/(\pi r^2))$$

per $r \ll R$ (risultato asintotico).

Differenze dalla percolazione discreta:

- La connettività dipende dalla geometria continua delle regioni
- Effetti di bordo rilevanti per sfere finite

- Densità critica dipende dalla forma delle regioni depositate

9.3 Percolazione di Livello per Campi Aleatori

Dato un campo aleatorio $h: S^2 \rightarrow \mathbb{R}$, definiamo l'**excursion set** al livello u :

$$A_u = \{p \in S^2 : h(p) \geq u\}$$

Domande fondamentali:

1. Qual è la topologia di A_u ? (numero di componenti connesse, genere, ...)
2. Come dipende dalla soglia u ?
3. Esiste u_c dove avviene una transizione di fase?

Teorema (Percolazione gaussiana): Per campo gaussiano isotropico h su S^2 con covarianza di lungo raggio, esiste livello critico $u_c \approx \mathbb{E}[h]$ (dipende dai dettagli della covarianza) tale che:

- $u < u_c$: A_u ha probabilità positiva di contenere un percorso che circonda completamente la sfera
- $u > u_c$: A_u consiste in isole disconnesse con alta probabilità

Formula di Euler-Poincaré aleatoria: Il numero di componenti connesse $C(u)$, handles $H(u)$, e buchi $B(u)$ soddisfano:

$$\mathbb{E}[C(u) - H(u) - B(u)] = \chi(S^2) = 2$$

dove χ è la caratteristica di Euler. Questo vincola la topologia media degli excursion sets.

Applicazione geofisica: Per campi di elevazione topografica, gli excursion sets A_u rappresentano regioni sopra quota u . La loro connettività determina:

- Bacini idrografici isolati
- Reti di drenaggio
- Accessibilità geografica

9.4 Funzioni di Correlazione e Critical Exponents

Vicino alla soglia critica, le proprietà di percolazione seguono leggi di potenza universali.

Probabilità di percolazione: Per p vicino a p_c :

$$P_\infty(p) = \mathbb{P}(\text{il cluster contenente l'origine è infinito}) \sim (p - p_c)^\beta$$

per $p > p_c$, dove β è l'**esponente critico** della probabilità di percolazione.

Lunghezza di correlazione: La scala tipica su cui le correlazioni decadono:

$$\xi(p) \sim |p - p_c|^{-\nu}$$

dove ν è l'esponente della lunghezza di correlazione.

Dimensione del cluster critico: Al punto critico $p = p_c$, la massa del cluster ha distribuzione:

$$n_s \sim s^{-\tau}$$

dove s è la dimensione del cluster e $\tau \approx 187/91 \approx 2.05$ per percolazione 2D.

Valori universali in 2D: Per reticoli bidimensionali, la teoria conforme (CFT) predice:

- $\beta = 5/36 \approx 0.139$
- $\nu = 4/3 \approx 1.333$
- $\gamma = 43/18 \approx 2.389$ (esponente della suscettività)

Su S^2 , la curvatura positiva modifica leggermente questi esponenti per sistemi finiti, ma nel limite termodinamico ($R \rightarrow \infty$) si recuperano i valori planari.

9.5 Percolazione Direzionale e Anisotropia

Per modellare reti di trasporto con direzioni preferenziali (catene montuose, valli fluviali), consideriamo **percolazione anisotropica**.

Modello: Su un reticolo sulla sfera, archi orizzontali (est-ovest) sono aperti con probabilità p_h , archi verticali (nord-sud) con probabilità p_v .

Diagramma di fase: Nel piano (p_h, p_v) , esistono regioni:

- **Fase disconnessa:** $\max(p_h, p_v) < \text{soglia critica}$
- **Fase connessa orizzontalmente:** p_h elevato, p_v basso
- **Fase connessa verticalmente:** p_v elevato, p_h basso
- **Fase totalmente connessa:** entrambi elevati

Anisotropia dei cluster: Per forte anisotropia ($p_h \gg p_v$), i cluster percolanti sono allungati orizzontalmente, con rapporto di aspetto:

$$\text{aspect ratio} \sim (p_h/p_v)^\zeta$$

dove $\zeta \approx 1/2$ è determinato dalla teoria critica.

Applicazione cartografica: Modellare reti stradali che seguono valli (direzione preferenziale) vs percorsi montani (direzione ostile). La percolazione anisotropica quantifica l'accessibilità in funzione della topografia direzionale.

9.6 Foreste Aleatorie e Spanning Trees su Grafi Planari

Una **foresta spanning** di un grafo G è un sottografo aciclico che connette tutti i vertici. La distribuzione uniforme su tutte le spanning trees genera un modello aleatorio con applicazioni a reti e ottimizzazione.

Algoritmo di Wilson: Un metodo elegante per campionare spanning tree uniforme su G :

1. Scegliere radice $r \in V$ arbitraria
2. Per ogni $v \neq r$:
 - Eseguire cammino aleatorio loop-erased da v fino a raggiungere la componente già connessa
 - Aggiungere questo percorso all'albero
3. Iterare fino a connettere tutti i vertici

Il risultato è una spanning tree con distribuzione uniforme.

Teorema di Kirchhoff: Il numero di spanning trees di G è dato dal determinante:

$$N_{\text{trees}}(G) = \det(Q_{\text{reduced}})$$

dove Q è la matrice Laplaciana del grafo:

- $Q = D - A$ (D matrice dei gradi, A matrice di adiacenza)
- Q_{reduced} : rimuovi una riga e colonna arbitrarie

Altezza della spanning tree aleatoria: Su grafi planari con n vertici, l'altezza (distanza massima dalla radice) di una spanning tree uniforme è:

$$\mathbb{E}[h_{\text{max}}] \sim c \sqrt{n}$$

con c dipendente dalla geometria del grafo.

9.7 Applicazione: Modellazione di Reti di Drenaggio Fluviale

I bacini fluviali formano naturalmente strutture ad albero: ogni punto sulla superficie drena verso un unico punto a valle, fino al mare.

Modello OCN (Optimal Channel Network): Generare reti di drenaggio minimizzando energia di dissipazione:

$$E[T] = \sum_{\{\text{canali } i\}} L_i^\alpha A_i^\beta$$

dove:

- T è la topologia della rete (spanning tree sul grafo topografico)
- L_i è la lunghezza del canale i
- A_i è l'area drenata a monte
- $\alpha \approx 1$, $\beta \approx 0.5$ sono esponenti empirici (scaling di Hack)

La rete ottimale minimizza $E[T]$ compatibilmente con vincoli topografici (elevazione decrescente verso valle).

Procedura di simulazione:

1. Partire da DEM (Digital Elevation Model) sulla regione
2. Costruire grafo diretto con archi da ogni cella verso le 8 celle adiacenti più basse
3. Applicare simulated annealing per trovare spanning tree di energia minima

4. Comparare con reti osservate

Leggi di scala osservate:

Le reti OCN riproducono:

- **Legge di Hack:** $L \sim A^h$ con $h \approx 0.57$ (lunghezza vs area drenata)
- **Distribuzione di Horton:** rapporto di biforcazione $R_B = N_k/N_{k+1} \approx 3-5$

dove N_k è il numero di streams di ordine k (classificazione di Strahler).

9.8 Minimum Spanning Forests e Clustering Geografico

Data una collezione di punti $\{p_i\}$ sulla sfera con pesi w_i , vogliamo partizionarli in cluster geograficamente coerenti.

MSF (Minimum Spanning Forest): Costruire Minimum Spanning Tree sull'intero insieme, poi rimuovere gli archi più lunghi fino a ottenere k componenti connesse.

Algoritmo:

1. Calcolare distanze geodetiche $d(p_i, p_j)$ per tutte le coppie
2. Costruire MST usando algoritmo di Kruskal o Prim:
 - Ordinare archi per distanza crescente
 - Aggiungere archi che non creano cicli
3. Rimuovere i $k-1$ archi più lunghi
4. Le k componenti connesse risultanti sono i cluster

Complessità: $O(n^2 \log n)$ per calcolo distanze + $O(E \log E)$ per MST, dove $E \leq n^2$ è il numero di archi.

Variante pesata: Se i punti hanno importanza variabile w_i (ad es., popolazione di città), usare distanza effettiva:

$$d_{\text{eff}}(i, j) = d(i, j) / (w_i w_j)^\alpha$$

con $\alpha > 0$ parametro che controlla l'influenza dei pesi.